



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 188896

TO: Anne-Marie ~~Baker~~ *Falk*  
Location: rem/2A45/2C18  
Art Unit: 1632  
Thursday, March 09, 2006  
Case Serial Number: 10/607712

From: Toby Port  
Location: Biotech-Chem Library  
REM-1A59  
Phone: 571-272-2523  
  
toby.port@uspto.gov

### Search Notes

Examiner ~~Baker~~, *Falk*.

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port  
X22523

*Please note - the pending files were not searched.  
If you need these please contact me.*

*Toby*

**THIS PAGE BLANK (USPTO)**



## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Anne-Marie Falk Examiner #: 74977 Date: 2/28/06  
 Art Unit: 1632 Phone Number: 20728 Serial Number: 10/607,712  
 Mail Box and Bldg/Room Location: L-2C18 L-2A45 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need. ME

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Human Methionine Synthase

Inventors (please provide full names): Roy A. Gravel

Earliest Priority Filing Date: 11/27/96

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

- Please search SEQ ID NO:1 (3919 nucleotides)

and

NA 3919

- also a size limited search to cover probes of

at least 18, 40, 50, or 60 nucleotides. (see attached claims)

(Although the claims cover a nucleic acid molecule of any length, I would like to see hits that are

100 nucleotides or smaller, in addition to hits that cover the full-length sequence).

- Please search commercial patents and pre-grant pubs databases.

Thanks!

## STAFF USE ONLY

## Type of Search

## Vendors and cost where applicable

Searcher: Per NA Sequence (#) \_\_\_\_\_ STN \_\_\_\_\_  
 Searcher Phone #: \_\_\_\_\_ AA Sequence (#) \_\_\_\_\_ Dialog \_\_\_\_\_  
 Searcher Location: \_\_\_\_\_ Structure (#) \_\_\_\_\_ Questel/Orbit \_\_\_\_\_  
 Date Searcher Picked Up: 3/6 Bibliographic \_\_\_\_\_ Dr.Link \_\_\_\_\_  
 Date Completed: 3/9 Litigation \_\_\_\_\_ Lexis/Nexis \_\_\_\_\_  
 Searcher Prep & Review Time: 8 Fulltext \_\_\_\_\_ Sequence Systems CS  
 Clerical Prep Time: \_\_\_\_\_ Patent Family \_\_\_\_\_ WWW/Internet \_\_\_\_\_  
 On the Time: 10 Other \_\_\_\_\_ Other (specify) \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

## Claims

1. A substantially pure human nucleic acid comprising at least 40 nucleotides that  
5 hybridizes under high stringency conditions to a sequence found within the nucleic acid  
of SEQ ID NO:1.

2. The nucleic acid of claim 1, wherein said sequence has a sequence  
complementary to at least 50% of at least 60 contiguous nucleotides of the nucleic acid  
10 encoding the methionine synthase polypeptide, said sequence sufficient to allow nucleic  
acid hybridization under high stringency conditions.

3. The nucleic acid of claim 1, wherein said nucleic acid comprises a mutation or  
a polymorphism, wherein said nucleic acid probe detects a mutation or polymorphism  
15 selected from the group consisting of D919G, H920D, and  $\Delta$ Ile881.

4. The nucleic acid of claim 3, wherein said sequence of said nucleic acid  
comprises the cobalamin binding domain of the human methionine synthase gene.

20 5. The nucleic acid of claim 2, wherein at least 18 contiguous nucleotides of said  
sequence are complementary to at least 90% of the corresponding nucleotides of the  
nucleic acid encoding the methionine synthase polypeptide.

6. The nucleic acid of claim 1, wherein said high stringency conditions comprise  
25 hybridization in 2X SSC at 40°C.

7. A substantially pure human nucleic acid, wherein the sequence of said nucleic  
acid is at least 75% identical to the corresponding region of at least 50 contiguous base  
pairs of the nucleic acid of SEQ ID NO:1.

**THIS PAGE BLANK (USPTO)**

8. A substantially pure human nucleic acid, wherein the sequence of said nucleic acid is at least 35% identical to the corresponding region of at least 50 contiguous base pairs of the nucleic acid of SEQ ID NO:1.

5

9. A kit for the analysis of a human methionine synthase nucleic acid, said kit comprising a nucleic acid probe useful for detecting in the nucleic acids of a human a mutation or polymorphism in said methionine synthase nucleic acid, wherein said mutation or polymorphism is selected from the group consisting of D919G, H920D, and  $\Delta$ Ile881.

10

10. The kit of claim 9, wherein said probe comprises at least 40 nucleotides that hybridizes at high stringency to a sequence found within the nucleic acid of SEQ ID NO:1.

15

**THIS PAGE BLANK** (USPTO,

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3919	100.0	3919	6	AR482278	AR482278	Sequence
2	3894.4	99.4	3894	6	HSU71285	U71285	Human 5-met
3	3884.8	99.1	7122	6	AR144957	AR144957	Sequence
4	3884.8	99.1	7122	6	AR367907	AR367907	Sequence
5	3884.8	99.1	7122	6	AX050442	AX050442	Sequence
6	3884.8	99.1	7122	6	AX069340	AX069340	Sequence
7	3884.8	99.1	7122	8	HSU75743	U75743	Human methi
8	3883.2	99.1	7224	6	AR367908	AR367908	Sequence
9	3883.2	99.1	7224	6	AX069342	AX069342	Sequence
10	3883.2	99.1	7224	8	HSU73338	U73338	Human methi
11	3876.2	98.9	7224	6	AR300095	AR300095	Sequence
12	3876.2	98.9	7224	6	AR438495	AR438495	Sequence
13	3876.2	98.9	7224	6	AR562003	AR562003	Sequence
14	3851	98.3	3856	6	AR482230	AR482230	Sequence
15	3785.2	96.6	3798	6	AQ930571	AQ930571	Sequence
16	3779.6	81.1	3832	4	DQ084519	DQ084519	Bos tauru
17	3669.4	78.3	4498	9	AF034214	AF034214	Rattus na
18	2417	61.7	5168	5	AJ720857	AJ720857	Gallus ga

Db 181 ATGATCCAGCGGAGAGCTAAACGAAGAACACATTCGAGGTCAGGAATTTAAAGATCAT 240  
Qy 241 GCCAGCCGCTGAAGGCAACATGACATTTTAAAGTATAATCTACGCTGATGTCATTTAC 300  
Db 241 GCCAGCCGCTGAAGGCAACATGACATTTTAAAGTATAATCTACGCTGATGTCATTTAC 300  
Qy 301 CAATATCCATAAGGAATATCTGCTGGCTGGGCGAGATATCATTTGAAAACAAATACATTTAGC 360  
Db 301 CAATATCCATAAGGAATATCTGCTGGCTGGGCGAGATATCATTTGAAAACAAATACATTTAGC 360  
Qy 361 AGCACTAGTATGCCCCAAGCTGATATGGCTTTGAACAATTTGGCTACCGGATGAACATG 420  
Db 361 AGCACTAGTATGCCCCAAGCTGATATGGCTTTGAACAATTTGGCTACCGGATGAACATG 420  
Qy 421 TGCTCTGAGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAG 480  
Db 421 TGCTCTGAGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAG 480  
Qy 481 AGGTTTGTGGCAGGGGCTCTGGGTCGGAATAATAGACACTCTCTGTGTCGCCATCTGTG 540  
Db 481 AGGTTTGTGGCAGGGGCTCTGGGTCGGAATAATAGACACTCTCTGTGTCGCCATCTGTG 540  
Qy 541 GAAAGCCGGATATAGGAAATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAG 600  
Db 541 GAAAGCCGGATATAGGAAATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAG 600  
Qy 601 GCCAAGGACTTCTGGATGGCGGGTGGATATCTTACTCATTTGAAACTATTTTGTATCT 660  
Db 601 GCCAAGGACTTCTGGATGGCGGGTGGATATCTTACTCATTTGAAACTATTTTGTATCT 660  
Qy 661 GCCAATGCCAAGGAGCCTGTTTGTCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
Db 661 GCCAATGCCAAGGAGCCTGTTTGTCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
Qy 721 CGGCTATCTTTATTCAGGAGCATCGTTGATATAAGTGGCGGACTCTTTCCGAGCAG 780  
Db 721 CGGCTATCTTTATTCAGGAGCATCGTTGATATAAGTGGCGGACTCTTTCCGAGCAG 780  
Qy 781 ACAGGAGAGGATTTGTCATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAAT 840  
Db 781 ACAGGAGAGGATTTGTCATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAAT 840  
Qy 841 TGTGCTTTGGTGAGCTGAGATGAGACTTTTATTTGAAATTAATTTGAAATATGTAACA 900  
Db 841 TGTGCTTTGGTGAGCTGAGATGAGACTTTTATTTGAAATTAATTTGAAATATGTAACA 900  
Qy 901 GCCTATGTCCTGTGTTATCCCAATGCAAGTCTTCCCAACACTTTTGGTGACTATGATGAA 960  
Db 901 GCCTATGTCCTGTGTTATCCCAATGCAAGTCTTCCCAACACTTTTGGTGACTATGATGAA 960  
Qy 961 AGCCCTTCTATGATGGCCAGCACCTAAAGATTTTGTATGGATGGCTTGGTCAATATA 1020  
Db 961 AGCCCTTCTATGATGGCCAGCACCTAAAGATTTTGTATGGATGGCTTGGTCAATATA 1020  
Qy 1021 GTTGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
Db 1021 GTTGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
Qy 1081 AATTTGAAGCCCTAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
Db 1081 AATTTGAAGCCCTAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
Qy 1141 CTAGAGCCCTCAGGATTTGAGCCGTACACCAACTTTTGTATTTGATTTGAGAGCGCTGTAAT 1200  
Db 1141 CTAGAGCCCTCAGGATTTGAGCCGTACACCAACTTTTGTATTTGATTTGAGAGCGCTGTAAT 1200  
Qy 1201 GTTGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1260  
Db 1201 GTTGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1260  
Qy 1261 TGTGTTGCCAAAGTGCAGGTGGAATGGAGCCAGGTTGTGGATGTCAACATGGATGAT 1320  
Db 1261 TGTGTTGCCAAAGTGCAGGTGGAATGGAGCCAGGTTGTGGATGTCAACATGGATGAT 1320

Qy 1321 GGCATGCTAGATGTCAGAGTGCAAATGACAGATTTTGGCAACTTAAATTTGCTTCCGAGCCA 1380  
Db 1321 GGCATGCTAGATGTCAGAGTGCAAATGACAGATTTTGGCAACTTAAATTTGCTTCCGAGCCA 1380  
Qy 1381 GACATCGCAAAAGGTACCTTTTGTGCATCGACTCTCTCCAAATTTTGTGTGATGGAAGCTGGG 1440  
Db 1381 GACATCGCAAAAGGTACCTTTTGTGCATCGACTCTCTCCAAATTTTGTGTGATGGAAGCTGGG 1440  
Qy 1441 TTTAAAGTGTCTGCCAAGGGAAGTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAC 1500  
Db 1441 TTTAAAGTGTCTGCCAAGGGAAGTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAC 1500  
Qy 1501 GACTCTTTGGAGAGGCGCAGGAAGATTTAAAGATGATGGAGTGTCTATGTTGGTGTATGGCT 1560  
Db 1501 GACTCTTTGGAGAGGCGCAGGAAGATTTAAAGATGATGGAGTGTCTATGTTGGTGTATGGCT 1560  
Qy 1561 TTTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCC 1620  
Db 1561 TTTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCC 1620  
Qy 1621 TACCATCTGCTTTGAAAAAACTGGGCTTTAAATCCAAATGACATTTTGTGACCCCTAAT 1680  
Db 1621 TACCATCTGCTTTGAAAAAACTGGGCTTTAAATCCAAATGACATTTTGTGACCCCTAAT 1680  
Qy 1681 ATCTTAACCATTTGGGACTTGGAAATGGAGGAAACAACTTTGATGCCATTTAAATTTATCCAT 1740  
Db 1681 ATCTTAACCATTTGGGACTTGGAAATGGAGGAAACAACTTTGATGCCATTTAAATTTATCCAT 1740  
Qy 1741 GCAACAAAAGTCAATTAAGAAACATTTACTGGAGCCAGAAATAAGTGGAGGCTTTTCCAAAC 1800  
Db 1741 GCAACAAAAGTCAATTAAGAAACATTTACTGGAGCCAGAAATAAGTGGAGGCTTTTCCAAAC 1800  
Qy 1801 TTGTCCTTCTCTTCCAGGAAATGGAAAGCCATTTGAGAGCAATGCAATGGGGTTTTCCTT 1860  
Db 1801 TTGTCCTTCTCTTCCAGGAAATGGAAAGCCATTTGAGAGCAATGCAATGGGGTTTTCCTT 1860  
Qy 1861 TACCATCAATCAAGTCTGGCATGGATGAGATAGTGAATGCTGGAACCTCCCTGTG 1920  
Db 1861 TACCATCAATCAAGTCTGGCATGGATGAGATAGTGAATGCTGGAACCTCCCTGTG 1920  
Qy 1921 TATGATGATATCCATAAGGAACTTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGAC 1980  
Db 1921 TATGATGATATCCATAAGGAACTTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGAC 1980  
Qy 1981 CCTGAGGCCACTGAGAAAGCTTTAGTTATGCCAGACTCAAGGCACAGGAGGGAAGAA 2040  
Db 1981 CCTGAGGCCACTGAGAAAGCTTTAGTTATGCCAGACTCAAGGCACAGGAGGGAAGAA 2040  
Qy 2041 GTCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAACGCTTTGAGTATGCCCTT 2100  
Db 2041 GTCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAACGCTTTGAGTATGCCCTT 2100  
Qy 2101 GTGAAGGCGCTTGAAGAAACATTTATTTAGGATATCTGAGGAAGCCAGGTTAAACCAAAA 2160  
Db 2101 GTGAAGGCGCTTGAAGAAACATTTATTTAGGATATCTGAGGAAGCCAGGTTAAACCAAAA 2160  
Qy 2161 AAATATCCCGAAGCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTT 2220  
Db 2161 AAATATCCCGAAGCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTT 2220  
Qy 2221 GGTGATCTTTTGGAGCTGGAAATGTTTCTACTCAGGTTTATAAAGTCAGCCGGGTT 2280  
Db 2221 GGTGATCTTTTGGAGCTGGAAATGTTTCTACTCAGGTTTATAAAGTCAGCCGGGTT 2280  
Qy 2281 ATGAAGAGGCTGTTGGCCACTTTATCCCTTTTATGAAAAAGAAAGAGAAACCCAGA 2340  
Db 2281 ATGAAGAGGCTGTTGGCCACTTTATCCCTTTTATGAAAAAGAAAGAGAAACCCAGA 2340  
Qy 2341 GTGCTTAACCGCACAGTAGAAGAAAGAGGACCCCTTACAGGGCACCATCTGCTGCGCCACT 2400  
Db 2341 GTGCTTAACCGCACAGTAGAAGAAAGAGGACCCCTTACAGGGCACCATCTGCTGCGCCACT 2400



2401 GTTAAAGCGACGTCGACGACATAGGCAAGCAACATAGTTGGAGTAGTCTTGGCTGCAAT 2460  
Db GTTAAAGCGACGTCGACGACATAGGCAAGCAACATAGTTGGAGTAGTCTTGGCTGCAAT 2460  
2461 AATTTCCGAGTTATGATTTAGGAGTCATGACTCCATGTGATAAGATGACTGAAAGCTGCT 2520  
Qy AATTTCCGAGTTATGATTTAGGAGTCATGACTCCATGTGATAAGATGACTGAAAGCTGCT 2520  
Db AATTTCCGAGTTATGATTTAGGAGTCATGACTCCATGTGATAAGATGACTGAAAGCTGCT 2520  
2521 CTTGACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTTCCCTCGATGAA 2580  
Qy CTTGACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTTCCCTCGATGAA 2580  
Db CTTGACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTTCCCTCGATGAA 2580  
2581 ATGATTTTGTGTCACAGGAATGGAGAGATTAGCTATAAGGATTCATTTGTTAGTGA 2640  
Qy ATGATTTTGTGTCACAGGAATGGAGAGATTAGCTATAAGGATTCATTTGTTAGTGA 2640  
Db ATGATTTTGTGTCACAGGAATGGAGAGATTAGCTATAAGGATTCATTTGTTAGTGA 2640  
2641 GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATATAGTCTCCAGATACAGTGCACCT 2700  
Qy GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATATAGTCTCCAGATACAGTGCACCT 2700  
Db GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATATAGTCTCCAGATACAGTGCACCT 2700  
2701 GTAATCCATGTCCTGGACGCTCCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAA 2760  
Qy GTAATCCATGTCCTGGACGCTCCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAA 2760  
Db GTAATCCATGTCCTGGACGCTCCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAA 2760  
2761 AATCTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGAC 2820  
Qy AATCTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGAC 2820  
Db AATCTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGAC 2820  
2821 CATTATGAGTCTCTCAAGGAGGAGATACATTTACCTTAACTCAAGCCAGAAAAGTGGT 2880  
Qy CATTATGAGTCTCTCAAGGAGGAGATACATTTACCTTAACTCAAGCCAGAAAAGTGGT 2880  
Db CATTATGAGTCTCTCAAGGAGGAGATACATTTACCTTAACTCAAGCCAGAAAAGTGGT 2880  
2881 TTCAAATGAGTGTGCTGTGAACCTCACCCAGTGAAGCCACGTTTATTTGGACCCAG 2940  
Qy TTCAAATGAGTGTGCTGTGAACCTCACCCAGTGAAGCCACGTTTATTTGGACCCAG 2940  
Db TTCAAATGAGTGTGCTGTGAACCTCACCCAGTGAAGCCACGTTTATTTGGACCCAG 2940  
2941 GTCCTTGAAGACTATGACTGAGAGCTGCTGAGGACTACATTTGAGTGAAGCCCTTCTTT 3000  
Qy GTCCTTGAAGACTATGACTGAGAGCTGCTGAGGACTACATTTGAGTGAAGCCCTTCTTT 3000  
Db GTCCTTGAAGACTATGACTGAGAGCTGCTGAGGACTACATTTGAGTGAAGCCCTTCTTT 3000  
3001 GATGCTGGAGCTCCGGGCAAGTACCCGATCGAGGCTTCCCAAGATATTTAAAGAC 3060  
Qy GATGCTGGAGCTCCGGGCAAGTACCCGATCGAGGCTTCCCAAGATATTTAAAGAC 3060  
Db GATGCTGGAGCTCCGGGCAAGTACCCGATCGAGGCTTCCCAAGATATTTAAAGAC 3060  
3061 AAAACAGTAGTGGAGAGCCAGAGGCTACATGATGATGCCAATATGCTGAAACACA 3120  
Qy AAAACAGTAGTGGAGAGCCAGAGGCTACATGATGATGCCAATATGCTGAAACACA 3120  
Db AAAACAGTAGTGGAGAGCCAGAGGCTACATGATGATGCCAATATGCTGAAACACA 3120  
3121 CTGATTAGTCAAAAGAAACTCCGGCCCGGGTGTGGTGTGTTGCTGCCAGCAGAGT 3180  
Qy CTGATTAGTCAAAAGAAACTCCGGCCCGGGTGTGGTGTGTTGCTGCCAGCAGAGT 3180  
Db CTGATTAGTCAAAAGAAACTCCGGCCCGGGTGTGGTGTGTTGCTGCCAGCAGAGT 3180  
3181 ATCCAGACGATTCACCTGTACGACAGCTGCTGTGCCCGAGGCTGACAGCCCAT 3240  
Qy ATCCAGACGATTCACCTGTACGACAGCTGCTGTGCCCGAGGCTGACAGCCCAT 3240  
Db ATCCAGACGATTCACCTGTACGACAGCTGCTGTGCCCGAGGCTGACAGCCCAT 3240  
3241 GCCACTTCTATGGTTAAGCAACAGCTCAGAGGACTGCTGCCAGCAGCAGCCTATAC 3300  
Qy GCCACTTCTATGGTTAAGCAACAGCTCAGAGGACTGCTGCCAGCAGCAGCCTATAC 3300  
Db GCCACTTCTATGGTTAAGCAACAGCTCAGAGGACTGCTGCCAGCAGCAGCCTATAC 3300  
3301 TACTGCTCTCAGACTTCAATGCTCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
Qy TACTGCTCTCAGACTTCAATGCTCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
Db TACTGCTCTCAGACTTCAATGCTCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3361 TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
Qy TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
Db TTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3421 GACTACAGCAGATCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTCAGAA 3480  
Qy GACTACAGCAGATCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTCAGAA 3480  
Db GACTACAGCAGATCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTCAGAA 3480  
3481 GAGCTCCATGAAAGAGTTCCGCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGAC 3540

3481 GAGCTCCATGAAAGAGTTCCCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGAC 3540  
Qy GAGCTCCATGAAAGAGTTCCCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGAC 3540  
3541 GTGCGAGACCTGCGAAGTTGCGGTACAAGGGCATCGCCGGCTCTGCTGCTACCCAGC 3600  
Db GTGCGAGACCTGCGAAGTTGCGGTACAAGGGCATCGCCGGCTCTGCTGCTACCCAGC 3600  
3541 GTGCGAGACCTGCGAAGTTGCGGTACAAGGGCATCGCCGGCTCTGCTGCTACCCAGC 3600  
Qy GTGCGAGACCTGCGAAGTTGCGGTACAAGGGCATCGCCGGCTCTGCTGCTACCCAGC 3600  
3601 CAGCCGACACACACCGAAGCTCAACATGTGGAGACTCGAGACATCGAGCAGTCTACA 3660  
Db CAGCCGACACACACCGAAGCTCAACATGTGGAGACTCGAGACATCGAGCAGTCTACA 3660  
3601 CAGCCGACACACACCGAAGCTCAACATGTGGAGACTCGAGACATCGAGCAGTCTACA 3660  
Qy CAGCCGACACACACCGAAGCTCAACATGTGGAGACTCGAGACATCGAGCAGTCTACA 3660  
3661 GGCAATTAGGTTAAGCAATCAATAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3720  
Db GGCAATTAGGTTAAGCAATCAATAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3720  
3661 GGCAATTAGGTTAAGCAATCAATAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3720  
Qy GGCAATTAGGTTAAGCAATCAATAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3720  
3721 TTCTCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATTCAGTTGAG 3780  
Db TTCTCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATTCAGTTGAG 3780  
3721 TTCTCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATTCAGTTGAG 3780  
Qy TTCTCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATTCAGTTGAG 3780  
3781 GATTATGATGAGGAAGAAACATATCTGTGGCTGAGTTGAGAAATGCTTGGACCCATT 3840  
Db GATTATGATGAGGAAGAAACATATCTGTGGCTGAGTTGAGAAATGCTTGGACCCATT 3840  
3781 GATTATGATGAGGAAGAAACATATCTGTGGCTGAGTTGAGAAATGCTTGGACCCATT 3840  
Qy GATTATGATGAGGAAGAAACATATCTGTGGCTGAGTTGAGAAATGCTTGGACCCATT 3840  
3841 TTGGGATATGATACAGACTAACTTTTGTGTGGGGAAGATTTTGTGTGTGTGTGTGTGTGT 3900  
Db TTGGGATATGATACAGACTAACTTTTGTGTGGGGAAGATTTTGTGTGTGTGTGTGTGTGT 3900  
3841 TTGGGATATGATACAGACTAACTTTTGTGTGGGGAAGATTTTGTGTGTGTGTGTGTGTGT 3900  
Qy TTGGGATATGATACAGACTAACTTTTGTGTGGGGAAGATTTTGTGTGTGTGTGTGTGTGT 3900  
3901 TCAAGGAATATCAACCTAG 3919  
Db TCAAGGAATATCAACCTAG 3919

## RESULT 2

HSU71285

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

1. .3917

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="1"

/map="1q43"

64. .3861

/notes="methionine synthase; cobalamin binding protein;

cytosolic protein"

/codon\_start=1

/product="5-methyltetrahydrofolate-homocysteine

methyltransferase"

HSU71285 3917 bp mRNA linear PRI 03-APR-1997  
Human 5-methyltetrahydrofolate-homocysteine methyltransferase mRNA,  
complete cds.

U71285  
U71285.1 GI:1923220  
vitamin B12; cobalamin binding site; methionine synthase.  
Homo sapiens (human)

1 (bases 1 to 3917)  
Leclerc, D., Campeau, E., Goyette, P., Adjalla, C.E., Christensen, B.,  
Ross, M., Eydoux, P., Rosenblatt, D.S., Rozen, R. and Gravel, R.A.  
Human methionine synthase: cDNA cloning and identification of  
mutations in patients of the cblG complementation group of  
folate/cobalamin disorders

Hum. Mol. Genet. 5 (12), 1867-1874 (1996)  
8968737  
2. (bases 1 to 3917)  
Leclerc, D.  
Direct Submission

Submitted (19-SEP-1996) Human Genetics, McGill University, Montreal  
Children's Hospital - Research Institute, Place Toulon, Room 222,  
4660 Ste-Catherine West, Montreal H3Z 2Z3, Canada  
On Apr 3, 1997 this sequence version replaced gi:1731672.

Location/Qualifiers  
1. .3917  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="1"

/map="1q43"

64. .3861  
/notes="methionine synthase; cobalamin binding protein;  
cytosolic protein"

/codon\_start=1  
/product="5-methyltetrahydrofolate-homocysteine  
methyltransferase"







1562 TTGATGAGGAGCAGCGCAACAGACAGACACAAATCAGAGTGTGCACCGGCGCT 1621  
1785 TTGATGAGGAGCAGCGCAACAGACAGACACAAATCAGAGTGTGCACCGGCGCT 1844  
1622 ACCATCTGCTGTGAAAGAACTCGGCTTTAATCCAAATGACATTAATTTTGGACCTTAATA 1681  
1845 ACCATCTGCTGTGAAAGAACTCGGCTTTAATCCAAATGACATTAATTTTGGACCTTAATA 1904  
1682 TCCTAAACATTGGGACTGGAAATGGAGGAAACAACTTGTATGCCATTAAATTTATCCATG 1741  
1905 TCCTAAACATTGGGACTGGAAATGGAGGAAACAACTTGTATGCCATTAAATTTATCCATG 1964  
1742 CAACAAAGCTCATTAAGAAACATTACTCTGAGCCAGAAATGAGGAGCTCTTCCAACT 1801  
1965 CAACAAAGCTCATTAAGAAACATTACTCTGAGCCAGAAATGAGGAGCTCTTCCAACT 2024  
1802 TGTCTCTTCTCTCCGAGGAATGGAAGCAATTCGAGAAAGCAATGCATGGGCTTTTCCTTT 1861  
2025 TGTCTCTTCTCTCCGAGGAATGGAAGCAATTCGAGAAAGCAATGCATGGGCTTTTCCTTT 2084  
1862 ACCATGCAATCAAGTCTGGGATGGACATGGAGATGATGAAATGCTGGAAACCTCCTCTGTG 1921  
2085 ACCATGCAATCAAGTCTGGGATGGACATGGGATGATGAAATGCTGGAAACCTCCTCTGTG 2144  
1922 ATGATGATATCCAAAGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAAC 1981  
2145 ATGATGATATCCAAAGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAAC 2204  
1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGAGAAAG 2041  
2205 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGAGAAAG 2264  
2042 TCATTTCAGACTGATGAGTGGAGAAATGCCCTGTGCGAAGAACGCTTGGATGATGCCCTTG 2101  
2265 TCATTTCAGACTGATGAGTGGAGAAATGCCCTGTGCGAAGAACGCTTGGATGATGCCCTTG 2324  
2102 TGAAGGCCATGGAAGAACATTAATGAGGATCTGAGGAGCAGGTTAAACCAAAAA 2161  
2325 TGAAGGCCATGGAAGAACATTAATGAGGATCTGAGGAGCAGGTTAAACCAAAAA 2384  
2162 AATATCCCGACCTCTCAATATATGAAGGACCCCTGATGAATGGAATGAATTCGTTG 2221  
2385 AATATCCCGACCTCTCAATATATGAAGGACCCCTGATGAATGGAATGAATTCGTTG 2444  
2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCGGGTTA 2281  
2445 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCGGGTTA 2504  
2282 TGAAGAGGCTGTGCGCACCTTATCCCTTTCATGGAAGAAAGAAAGAACCAAGAG 2341  
2505 TGAAGAGGCTGTGCGCACCTTATCCCTTTCATGGAAGAAAGAAAGAACCAAGAG 2564  
2342 TGCTTTACGGCACAGTAGAAGAGGACCTTACAGGCGCACCATCGTGTGCGCACTG 2401  
2565 TGCTTTACGGCACAGTAGAAGAGGACCTTACAGGCGCACCATCGTGTGCGCACTG 2624  
2402 TTAAAGGCGAGTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2461  
2625 TTAAAGGCGAGTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2684  
2462 ATTTCCGAGTTATGATTTAGAGTCAATGACTCCTCAATGATAGATCTGAAGAGCTGCTC 2521  
2685 ATTTCCGAGTTATGATTTAGAGTCAATGACTCCTCAATGATAGATCTGAAGAGCTGCTC 2744  
2522 TTGACCAAGAGCAGATATATTTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA 2581  
2745 TTGACCAAGAGCAGATATATTTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA 2804  
2582 TGATTTTGTGCGCAAGGAAATGAGAGATGAGTCTATAAGGATTCATTTGTTGATGGAG 2641  
2805 TGATTTTGTGCGCAAGGAAATGAGAGATGAGTCTATAAGGATTCATTTGTTGATGGAG 2864  
2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701

2865 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACCTG 2924  
2702 TAATTCATGTCTCTGGAGCGGTCCAAGAGTGTGTGGTGTGTTCCAGCTCTTTAGATGAAA 2761  
2925 TAATTCATGTCTCTGGAGCGGTCCAAGAGTGTGTGGTGTGTTCCAGCTCTTTAGATGAAA 2984  
2762 ATCTAAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGACC 2821  
2985 ATCTAAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGACC 3044  
2822 ATTATGAGTCTCTCAAGGAGGAGATACTTACCTTAACTCAAGCCAGAAAAAGTGGTT 2881  
3045 ATTATGAGTCTCTCAAGGAGGAGATACTTACCTTAACTCAAGCCAGAAAAAGTGGTT 3104  
2882 TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGAGCCAGG 2941  
3105 TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGAGCCAGG 3164  
2942 TCTTTGAAGACTATGACCTGCAGAACTGTGTGACTACATTTGACTGGAAGCCTTTCTTTG 3001  
3165 TCTTTGAAGACTATGACCTGCAGAACTGTGTGACTACATTTGACTGGAAGCCTTTCTTTG 3224  
3002 ATGCTCTGGCAGCTCCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA 3061  
3225 ATGCTCTGGCAGCTCCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA 3284  
3062 AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCACAATATCTGGAACAC 3121  
3285 AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCACAATATCTGGAACAC 3344  
3122 TGATTAGTCAAAAAAACTCCGGGCGCGGGTGTGTGGTGTGTTGGCCACGACAGAGTA 3181  
3345 TGATTAGTCAAAAAAACTCCGGGCGCGGGTGTGTGGTGTGTTGGCCACGACAGAGTA 3404  
3182 TCCAAAGCAGCATTCACCTGTACGAGAGGCTCTGTGCCCGCCAGGCTGACAGGCCATAG 3241  
3405 TCCAAAGCAGCATTCACCTGTACGAGAGGCTCTGTGCCCGCCAGGCTGACAGGCCATAG 3464  
3242 CCACTTTCTATGGTTTAAAGCAACAGGCTGAGAGGACTCTGCGCAGCACCGAGCCATACT 3301  
3465 CCACTTTCTATGGTTTAAAGCAACAGGCTGAGAGGACTCTGCGCAGCACCGAGCCATACT 3524  
3302 ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCGCGCTACCTGAGCTGAGCTGCTGT 3361  
3525 ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCGCGCTACCTGAGCTGAGCTGCTGT 3584  
3362 TTGCCGTTGCTGCTTTGGGTAGAGAGCTGAGCAGGCTTATGAGGATGATGGTCAAG 3421  
3585 TTGCCGTTGCTGCTTTGGGTAGAGAGCTGAGCAGGCTTATGAGGATGATGGTCAAG 3644  
3422 ACTACAGCAGCATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTTGCAGAG 3481  
3645 ACTACAGCAGCATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTTGCAGAG 3704  
3482 AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGAGCTGAGAG 3541  
3705 AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGAGCTGAGAG 3764  
3542 TCCGAGACCTGCGAAGTTGGGTACAGGGCATCCCGCGCTCTGGCTACCCAGCC 3601  
3765 TCCGAGACCTGCGAAGTTGGGTACAGGGCATCCCGCGCTCTGGCTACCCAGCC 3824  
3602 AGCCCGACCAACCGAGAGGCTCACCATGTGGAGACTCGCAGACATCGACAGTCTACAG 3661  
3825 AGCCCGACCAACCGAGAGGCTCACCATGTGGAGACTCGCAGACATCGACAGTCTACAG 3884  
3662 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGCTTCCAGCAGTCTCAGGCTCTACT 3721  
3885 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGCTTCCAGCAGTCTCAGGCTCTACT 3944  
3722 TCTCCAAATTTGAAAGTCCAAATATTTTGTCTGTGGGGAAGATTTTCCAGGATCAGGTTGAG 3781





1785	TTGATGAGAAAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCAACCGGGCCCT	1844	
Qy	1622	ACCATCTGCTTGTGAAAAAACTGGGCTTTTAATCCAAATGACATTAATTTTTTGACCCCTAAATA	1681
	1845	ACCATCTGCTTGTGAAAAAACTGGGCTTTTAATCCAAATGACATTAATTTTTTGACCCCTAAATA	1904
	1682	TCCTAACCATTTGGGACTGGGAATGGAGGACACAACCTTGATATGCAATTAATTTTTATCCATG	1741
Db	1905	TCCTAACCATTTGGGACTGGGAATGGAGGAAACACAACCTTGATATGCAATTAATTTTTATCCATG	1964
	1742	CAACAAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACCT	1801
	1965	CAACAAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACCT	2024
Qy	1802	TGTCCTTCTCTCTCCGAGGAATGGAAGCCATTCCGAGAAGCAATTCGATGGGGTTTTTCCCTTT	1861
	2025	TGTCCTTCTCTCTCCGAGGAATGGAAGCCATTCCGAGAAGCAATTCGATGGGGTTTTTCCCTTT	2084
	1862	ACCATGCAATCAAGTCTGGCAATGGACATGAGAGATAGTGAATGCTTGGAAACCTCCCTGTGTG	1921
Db	2085	ACCATGCAATCAAGTCTGGCAATGGACATGGGGATAGTGAATGCTTGGAAACCTCCCTGTGTG	2144
	1922	ATGATGATATCCATAAGGAACCTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC	1981
	2145	ATGATGATATCCATAAGGAACCTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC	2204
Qy	1982	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCCAGACTCAAGGCACAGGAGGGAAGAAAG	2041
	2205	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCCAGACTCAAGGCACAGGAGGGAAGAAAG	2264
	2042	TCAATTCAGACTGATAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG	2101
Db	2265	TCAATTCAGACTGATAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG	2324
	2102	TGAAGGCCATTGAAAAACATATTATTGAGGATCTGAGGAGGCCAGGTGTAACCCAAAAA	2161
	2325	TGAAGGCCATTGAAAAACATATTATTGAGGATCTGAGGAGGCCAGGTGTAACCCAAAAA	2384
Qy	2162	AATATCCCCGACCTCTCAATATTAATCGAAGACCCCTGATGAATGGAATGAAATTTGTTG	2221
	2385	AATATCCCCGACCTCTCAATATTAATCGAAGACCCCTGATGAATGGAATTTGTTG	2444
	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAAAGTCAGCCCGGGTTA	2281
Db	2445	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAAAGTCAGCCCGGGTTA	2504
	2282	TGAAGAAGGCTGTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAGAACCCAGAG	2341
	2505	TGAAGAAGGCTGTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAGAACCCAGAG	2564
Qy	2342	TGCTTAAACGACACAGTAGAAGAGAGAGACCCCTTACAGGGCACCCTCGTCTGGCCACTG	2401
	2565	TGCTTAAACGACACAGTAGAAGAGAGAGACCCCTTACAGGGCACCCTCGTCTGGCCACTG	2624
	2402	TTAAAGGCGAGCTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTTGGCTGCAATA	2461
Db	2625	TTAAAGGCGAGCTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTTGGCTGCAATA	2684
	2462	ATTTCCGAGTTATTGATTTAGGAGTCATGACTCATGTGTATAGATTAAGTAAAGCTGCTC	2521
	2685	ATTTCCGAGTTATTGATTTAGGAGTCATGACTCATGTGTATAGATTAAGTAAAGCTGCTC	2744
Qy	2522	TTGACCAACAAAGCAGATATTAATTTGGCCCTGTGAGGACCTCATCACCTCCCTCCCTGATGAAA	2581
	2745	TTGACCAACAAAGCAGATATTAATTTGGCCCTGTGAGGACCTCATCACCTCCCTCCCTGATGAAA	2804
	2582	TGATTTTTTGTTCGCAAGGAATGAGAGATTAGTCTATAAGGATTTCCATTGTTGATTTGGAG	2641
Db	2805	TGATTTTTTGTTCGCAAGGAATGAGAGATTAGTCTATAAGGATTTCCATTGTTGATTTGGAG	2864
	2642	GAGCAACCACTTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCAACCTG	2701
	2865	GAGCAACCACTTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCAACCTG	2924

QY		2702	TAAATCCATGTCCTGGACCGGTCCAAGAGTGTGTGTGTGTGCCAGTGTTTAGATGAATAA	2767
DB		2925	TAAATCCATGTCCTGGACCGGTCCAAGAGTGTGTGTGTGTGCCAGTGTTTAGATGAATAA	2984
QY		2762	ATCTAAAAGGATGAATACTTTTGAGGAAATCATGCAAGAATATGAAGATATTAGACAGGACC	2821
DB		2985	ATCTAAAAGGATGAATACTTTTGAGGAAATCATGAGAGNATATGAAGATATTAGACAGGACC	3044
QY		2822	ATTATAGTCTCTCAAAGGAGGAGATPACTTAACCTTTAAGTCAAGGCAGAAAAAGTGGTT	2881
DB		3045	ATTATAGTCTCTCAAAGGAGGAGATPACTTAACCTTTAAGTCAAGGCAGAAAAAGTGGTT	3104
QY		2882	TCCAAATGGATTTGGTCTGTGAACCTCACCAGTGAAGCCCCAGTTTATTTTGGAGACCCAGG	2941
DB		3105	TCCAAATGGATTTGGTCTGTGAACCTCACCAGTGAAGCCCCAGTTTATTTTGGAGACCCAGG	3164
QY		2942	TCCTTTGAAGACTATGACCTCCAGAAAGCTGTGTGACATACATTGACTGGAAGCCTTTCTTTG	3001
DB		3165	TCCTTTGAAGACTATGACCTCCAGAAAGCTGTGTGACATACATTGACTGGAAGCCTTTCTTTG	3224
QY		3002	ATGTCCTGGACGCTCCGGGGCAAGTACCOCGAATCGAGGCTTTCCCACAAGATATTAAACGACA	3061
DB		3225	ATGTCCTGGACGCTCCGGGGCAAGTACCOCGAATCGAGGCTTTCCCACAAGATATTAAACGACA	3284
QY		3062	AAACAGTAGTGGAGAGGCCAGGAAGTCTACAGATAGTCCCAAAATPATGCTGAACACAC	3121
DB		3285	AAACAGTAGTGGAGAGGCCAGGAAGTCTACAGATAGTCCCAAAATPATGCTGAACACAC	3344
QY		3122	TGATTAGTCAAAGAAANCTCGGGCCCGGGTGTGGTGTGGTCTGGCCAGCACAGAGTA	3181
DB		3345	TGATTAGTCAAAGAAANCTCGGGCCCGGGTGTGGTGTGGTCTGGCCAGCACAGAGTA	3404
QY		3182	TCCAAGACGACATTCACCTGTACGCAGAGCGTGTGTGCCCCAGGCTGCAGAGGCCCATAG	3241
DB		3405	TCCAAGACGACATTCACCTGTACGCAGAGCGTGTGTGCCCCAGGCTGCAGAGGCCCATAG	3464
QY		3242	CCACTTTCTATGGGTTAAGGCACACAGGCTGAGAAAGGACTGTGCCAGCACGGAGGCCATACT	3301
DB		3465	CCACTTTCTATGGGTTAAGGCACACAGGCTGAGAAAGGACTGTGCCAGCACGGAGGCCATACT	3524
QY		3302	ACTGCCCTCHAGACTTCATCGCTCCCTTGCAATCTTGGCANCTCGTGACTACCTGGGCCCTGT	3361
DB		3525	ACTGCCCTCHAGACTTCATCGCTCCCTTGCAATCTTGGCANCTCGTGACTACCTGGGCCCTGT	3584
QY		3362	TTGCCGTTGCTGTCTTTGGSGTAGAAGAGCTGAGCAAGGCTATGAGAGTATGAGTGTGACG	3421
DB		3585	TTGCCGTTGCTGTCTTTGGSGTAGAAGAGCTGAGCAAGGCTATGAGAGTATGAGTGTGACG	3644
QY		3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCAGAAG	3481
DB		3645	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCAGAAG	3704
QY		3482	AGCTCCATGAAAGATTCCGCCGAAACTGTGGGCCCTACTGTGGCAGTGAAGCAGCTGGACG	3541
DB		3705	AGCTCCATGAAAGATTCCGCCGAAACTGTGGGCCCTACTGTGGGCCCTACTGTGGCAGTGAAGCAGCTGGACG	3764
QY		3542	TCGCAGACCTGCGAAGTTCGGGTACAAGGGCATCCGCCCGGCTCTGGCTACCCCGAGCC	3601
DB		3765	TCGCAGACCTGCGAAGTTCGGGTACAAGGGCATCCGCCCGGCTCTGGCTACCCCGAGCC	3824
QY		3602	AGCCCGACCAACCCGAGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3661
DB		3825	AGCCCGACCAACCCGAGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3884
QY		3662	GCAATTAGGTTAACGAATCATTTAGCAATGGCACCTGCTTCAAGCAGTCTCAGGCTCTTACT	3721
DB		3885	GCAATTAGGTTAACGAATCATTTAGCAATGGCACCTGCTTCAAGCAGTCTCAGGCTCTTACT	3944
QY		3722	TCTCCAAATTTGAATGCCAAATATTTTTGCTGTGGGGAAGATTTTCAAGATCATAGTTGAGG	3781
DB		3945	TCTCCAAATTTGAATGCCAAATATTTTTGCTGTGGGGAAGATTTTCAAGATCATAGTTGAGG	4004







Db 3945 TCTCCAAATTGAAGTCCAAATATTTTGTCTGTGGGGAAGATTTCACAGGATCAGGTTGAGG 4004  
Qy 3782 ATTATGCAATGAGGAAGAACATATCTGTGGCTGTAGGTTGAGAAATGGCTTGGACCCATT 3841  
Db 4005 ATTATGCAATGAGGAAGAACATATCTGTGGCTGTAGGTTGAGAAATGGCTTGGACCCATT 4064  
Qy 3842 TGGGATATGATACAGACTAACTTTTGTCTGTGGGGAAGATTTCACAGGATCAGGTTGAGG 3901  
Db 4065 TGGGATATGATACAGACTAACTTTTGTCTGTGGGGAAGATTTCACAGGATCAGGTTGAGG 4122  
Qy 3902 CAAGGAATACACCTAG 3919  
Db 4123 CAAGGAATACACCTAG 4140

RESULT 6  
LOCUS AX069340 7122 bp DNA linear PAT 25-JAN-2001  
DEFINITION Sequence 4 from Patent WO0102600.  
ACCESSION AX069340  
VERSION AX069340.1 GI:12579203  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Yuan, C.S.  
TITLE Detection of analytes using attenuated enzymes  
JOURNAL Patent: WO 0102600-A 4 11-JAN-2001;  
GENERAL ATOMICS (US)

FEATURES  
source  
1..7122  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
287..4084  
/note="unnamed protein product; Human methionine synthase  
CDNA"

/codon\_start=1  
/protein\_id="CAC27299.1"  
/db\_xref="GI:12579204"  
/translation="MSPALQDLSOPEGLKLTDRBINAIIQKIMVLVDGGMGTIQRE  
KLNREHFGQFQKHARPLKGNNDLSITOPDVIQHKVEYLLAGADIENITFSSTS  
IAQADYGLHLYRNMWCSAGVARKAEEVLTGTGKFPVAGALGPTNKTLSVSPVE  
RPYRNIITFDELVEAYQQAQGLDGGVDIIILITFDANAKALFALQNLFEKYA  
PRPIFSGTIVDKSGRTLSGGTGEFVIVSHGELPYTGLNCALGAEMRPFIIIGK  
CTTAYLCVPNAGLPNTFGDYDETSPMAKHLKOPAMDGLVNIIVGGCGSPDHIREI  
AEAVNCKRPVPPAPAFRGMLLSLGEPFRIGPYTNFVNIIGERNVAGSRKPAKLIMA  
GNVEALCVAKVQVEMGAQVLDVNMDDGMLDGPASMTFCNLIASEPDIKVPICIDS  
SNFATEAGLKKCCQKCIINSISLKEGDDDFLEKARKIKKYGAAMVWAFBEGQATE  
TDTKIRVCTRAYHLVLLKGLGNPNNDIIIPDNILITGTGMEHNLYAINFIHATVKIE  
TLPGARIISGLSINLSFSFRGMEAREAMHGVFLYHAIKSGMDMGIVNAGNIPVDYIH  
KELQLCEDLWKNQPEATKLLRYAQGTGGKVIQTDWRNGPVEERLEYALVKG  
LEKHIIDTEARLNQKYRPLNIIIEGLPMNGKIVGLDFRGMFLPQVTKSARVM  
KKAUGHLLPPEKEEREZVLNGTVEEDPYQGTIVLATVKGDDHDIKGNIVGVLC  
NNFRVIDGVTMPCDKIIKALDHKADIIGLSGLITPSIDEMIIVAKEMERLAIPI  
LIGGATSKHTATVAPRYAPVHVLDAKSVVCSQLLDENLKDYPBEIMEEYE  
DIRQHYFELSKERRYLPISQARKSGFQMDLSEPHVPKPTFTIGTVQFBDYDLKLVDY  
IDWKFDPVWOLRGYKPNRFPKI FNDKTVGSEARKVYDDAHNMLNTLISOKLRLARG  
VYGFPAOSIIDDHLYAEAAVPOAEPFIATFYGLRQQAEEKDSATEPYKLSDFIAP  
LHSGTRDYLGIYPVACPFGEVLSKAYEDDDYSIMVKGALGDRLABEPSELHERVR  
RELWAYCSQSEQDLADRLRYKIRPAPFQSDPDYSDHTEKUTMRLADIESTGIRLT  
ESLAWAPASAVSLGYFSLNKSIFYAVGKISKDQVEDYALRKNIISVAREKVMILGTY  
DTD"

## ORIGIN

Query Match 99.1%; Score 3884.8; DB 6; Length 7122;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3904; Conservative 12; Indels 2; Gaps 1;  
Qy 2 GTCACTGTGGAGACGACGCTTCTCTGTGGCGGCCCTCTGCGAAGGAGGAGACTCGACA 61

Db 225 GTCACTGTGGAGACGACGCTTCTCTGTGGCGGCCCTCTGCGAAGGAGGAGACTCGACA 284  
Qy 62 ACATGTCCACCGCGCTCCAAAGACCTGTGCAACCCGAAAGGTCTGAAAGAAACCTCGCGG 121  
Db 285 ACATGTCCACCGCGCTCCAAAGACCTGTGCAACCCGAAAGGTCTGAAAGAAACCTCGCGG 344  
Qy 122 ATGAGATCAATGCCATTCTGCAAGAGAGGATTATGTGTCTGGATGGAGGATGGGACCA 181  
Db 345 ATGAGATCAATGCCATTCTGCAAGAGAGGATTATGTGTCTGGATGGAGGATGGGACCA 404  
Qy 182 TGATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTGAGGAATTTAAAGATCATG 241  
Db 405 TGATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTGAGGAATTTAAAGATCATG 464  
Qy 242 CCAGGCGCTGAAAGGCAACAATGACATTTTAACTATAAATCTCAGCCTGATGTCTATTACC 301  
Db 465 CCAGGCGCTGAAAGGCAACAATGACATTTTAACTATAAATCTCAGCCTGATGTCTATTACC 524  
Qy 302 AAATCCATAAGGAATACTTGTCTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA 361  
Db 525 AAATCCATAAGGAATACTTGTCTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA 584  
Qy 362 GCATAGTATTTGCCAAGCTGACTATGGCCTTGAAACAATTGGGCTACCGGATGAACATGT 421  
Db 585 GCATAGTATTTGCCAAGCTGACTATGGCCTTGAAACAATTGGGCTACCGGATGAACATGT 644  
Qy 422 GCTCTGAGGAGTGGCGAGAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 481  
Db 645 GCTCTGAGGAGTGGCGAGAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 704  
Qy 482 GGTTTGTGGCAGGGGCTCTGGGTCGACATAAAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 705 GGTTTGTGGCAGGGGCTCTGGGTCGACATAAAGACACTCTCTGTGTCCCATCTGTGG 764  
Qy 542 AAAGCGCGGATTAAGGAAACATCACTTGTGATGAGCTTTGTAAGACATACCAAGAGCAGG 601  
Db 765 AAAGCGCGGATTAAGGAAACATCACTTGTGATGAGCTTTGTAAGACATACCAAGAGCAGG 824  
Qy 602 CCAAGGACTCTCTGGATGGGGGTGATATCTTACTCATTTGAAACTATTTTTCATCTG 661  
Db 825 CCAAGGACTCTCTGGATGGGGGTGATATCTTACTCATTTGAAACTATTTTTCATCTG 884  
Qy 662 CCAATGCCAAGCAGCCTTGTGTGACCTCCAAATCTTTTGTAGGAGAAATATGCTCCC 721  
Db 885 CCAATGCCAAGCAGCCTTGTGTGACCTCCAAATCTTTTGTAGGAGAAATATGCTCCC 944  
Qy 722 GGCCTATCTTTATTTTCAAGGACGATCGTTGATAAAAGTGGCGGACCTCTTCCGACAGA 781  
Db 945 GGCCTATCTTTATTTTCAAGGACGATCGTTGATAAAAGTGGCGGACCTCTTCCGACAGA 1004  
Qy 782 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGGAGACCACTCTGCATTTGGATTAAAT 841  
Db 1005 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGGAGAACCACTCTACATTTGGATTAAAT 1064  
Qy 842 GTGCTTTGGGTGACGTGAGATGAGACTTTTATTTGAAATAAATTTGAAATAATTTACACAG 901  
Db 1065 GTGCTTTGGGTGACGTGAGATGAGACTTTTATTTGAAATAAATTTGAAATAATTTACACAG 1124  
Qy 902 CCTATGCTCTCTGTTATCCCAATCAGGTCTTCCCAACAACCTTTTGGTGACTATGATGAAA 961  
Db 1125 CCTATGCTCTCTGTTATCCCAATCAGGTCTTCCCAACAACCTTTTGGTGACTATGATGAAA 1184  
Qy 962 CGCCTTCTATGATGGCCAGCACCTTAAAGGATTTTGTCTATGGATGGCTTGGCTCAATAG 1021  
Db 1185 CGCCTTCTATGATGGCCAGCACCTTAAAGGATTTTGTCTATGGATGGCTTGGCTCAATAG 1244  
Qy 1022 TTGAGGAGTCTGTGGGTCAACACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAA 1081  
Db 1245 TTGAGGAGTCTGTGGGTCAACACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAA 1304  
Qy 1082 ATTGTAAAGCTTAGAGTTCCACTGCGCACTGCTTTTGAAGGACATATGTTACTGTCTGTC 1141

Db	1305	ATTGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTCTGCTC	1364
Qy	1142	TAGAGCCCTTCAGGATTGGACCGTACACCAACTTTGTTTAACTTTGGAGAGCGCTGTAAATG	1201
Db	1365	TAGAGCCCTTCAGGATTGGACCGTACACCAACTTTGTTTAACTTTGGAGAGCGCTGTAAATG	1424
Qy	1202	TTGCAGGATCAAGGAAGTTTCTAACTCATCATGCGCAGGAACCTATGAAAGACCTTGT	1261
Db	1425	TTGCAGGATCAAGGAAGTTTCTAACTCATCATGCGCAGGAACCTATGAAAGACCTTGT	1484
Qy	1262	GTGTTGCCAAAGTGCAGGTGGAATGGGAGCCAGGTGTTGGATGTCAACATGATGATG	1321
Db	1485	GTGTTGCCAAAGTGCAGGTGGAATGGGAGCCAGGTGTTGGATGTCAACATGATGATG	1544
Qy	1322	GCATGCTAGATGGTCCAAAGTCAATGACCAAGATTTTGCACCTTAATTTGCTCCAGCCAG	1381
Db	1545	GCATGCTAGATGGTCCAAAGTCAATGACCAAGATTTTGCACCTTAATTTGCTCCAGCCAG	1604
Qy	1382	ACATCGCAAGGTACTTTTGTGTCATCGACTCTCCAAATTTTGTGATGTAAGCTGGGT	1441
Db	1605	ACATCGCAAGGTACTTTTGTGTCATCGACTCTCCAAATTTTGTGATGTAAGCTGGGT	1664
Qy	1442	TAAAGTCTGCTCCAAAGGAAGTGCAATGATAGCAATTTAGTCTGAAGGAAGGAGGACG	1501
Db	1665	TAAAGTCTGCTCCAAAGGAAGTGCAATGATAGCAATTTAGTCTGAAGGAAGGAGGACG	1724
Qy	1502	ACTTCTTGGAGAGGCCAGGAAGATTTAAAGTAGTATGGAGCTGCTATGGTGTCTATGGCTT	1561
Db	1725	ACTTCTTGGAGAGGCCAGGAAGATTTAAAGTAGTATGGAGCTGCTATGGTGTCTATGGCTT	1784
Qy	1562	TTGATGAAGAGGACGAGCAACAGAAACAGACACAAATAATCAGATGTGCAACCCGGGCT	1621
Db	1785	TTGATGAAGAGGACGAGCAACAGAAACAGACACAAATAATCAGATGTGCAACCCGGGCT	1844
Qy	1622	ACCATCTGCTGTGAAAGAACTGGGCTTTTAAATGCAATTTATTTTTCGACCTTAATA	1681
Db	1845	ACCATCTGCTGTGAAAGAACTGGGCTTTTAAATGCAATTTATTTTTCGACCTTAATA	1904
Qy	1682	TCCTAACCATTTGGGACTGGATGGAGGAAACAACTTGTATGCCAATTAATTTTATCCATG	1741
Db	1905	TCCTAACCATTTGGGACTGGATGGAGGAAACAACTTGTATGCCAATTAATTTTATCCATG	1964
Qy	1742	CAACAAAGTCAATTAAGAAACATTTACTGAGGCCAGAAATAGTGGAGTCTTTTCCAACT	1801
Db	1965	CAACAAAGTCAATTAAGAAACATTTACTGAGGCCAGAAATAGTGGAGTCTTTTCCAACT	2024
Qy	1802	TGTCCTCTCTCCGAGGAATGAGGCAATTCGAGGCAATGCAATGGGCTTTTCCCTTT	1861
Db	2025	TGTCCTCTCTCCGAGGAATGAGGCAATTCGAGGCAATGCAATGGGCTTTTCCCTTT	2084
Qy	1862	ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCTCCTGTGT	1921
Db	2085	ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCTCCTGTGT	2144
Qy	1922	ATGATGATATCCATAAGGAACCTCTGACAGCTCTGTGAAGATCTCATCTGGAATAAGACC	1981
Db	2145	ATGATGATATCCATAAGGAACCTCTGACAGCTCTGTGAAGATCTCATCTGGAATAAGACC	2204
Qy	1982	CTGAGGCCACTGAGAGCTCTTAACCTTATGCCAGACTCAAGGCAAGGAGGGAAGAAAG	2041
Db	2205	CTGAGGCCACTGAGAGCTCTTAACCTTATGCCAGACTCAAGGCAAGGAGGGAAGAAAG	2264
Qy	2042	TCATTCACATCATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAGTATGCCCTTG	2101
Db	2265	TCATTCACATCATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAGTATGCCCTTG	2324
Qy	2102	TGAAGGCAATTTGAAAAACATATTTAGGAGTACTGAGGAAGCCAGGTAAACCAAAAAA	2161
Db	2325	TGAAGGCAATTTGAAAAACATATTTAGGAGTACTGAGGAAGCCAGGTAAACCAAAAAA	2384
Qy	2162	AATATCCCGACCTCTCAATTAATTTAGAGGACCCCTGATGAATGGAATGAATTTGTTG	2221
Db	2385	AATATCCCGACCTCTCAATTAATTTAGAGGACCCCTGATGAATGGAATGAATTTGTTG	2444

Qy	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTTATAAGTCAAGCCGGGTTA	2281
Db	2445	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTTATAAGTCAAGCCGGGTTA	2504
Qy	2282	TGAAGAGGCTGTGTCACCTTATCCCTTTTCAATGAAAAAAGAAAGAAACCAAGAG	2341
Db	2505	TGAAGAGGCTGTGTCACCTTATCCCTTTTCAATGAAAAAAGAAAGAAACCAAGAG	2564
Qy	2342	TGCTTAAACGACAGTAGTAAGAGAGGACCTTTACCAAGGCAACATCGTCTGCGCACTG	2401
Db	2565	TGCTTAAACGACAGTAGTAAGAGAGGACCTTTACCAAGGCAACATCGTCTGCGCACTG	2624
Qy	2402	TTAAAGGCGAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTTGGCTGCAATA	2461
Db	2625	TTAAAGGCGAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTTGGCTGCAATA	2684
Qy	2462	ATTTCGAGTATTGATTTTAGGAGTCAATGACTCCATGTAAGATACCTGAAAGCTGCTC	2521
Db	2685	ATTTCGAGTATTGATTTTAGGAGTCAATGACTCCATGTAAGATACCTGAAAGCTGCTC	2744
Qy	2522	TTGACCAAAAGCAGATATAATTTGGCCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2581
Db	2745	TTGACCAAAAGCAGATATAATTTGGCCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2804
Qy	2582	TGATTTTGTGCAAGAAATGGAGAGATTAAGTATAAGATTTCCATTTGTTGATGGAG	2641
Db	2805	TGATTTTGTGCAAGAAATGGAGAGATTAAGTATAAGATTTCCATTTGTTGATGGAG	2864
Qy	2642	GAGCAACCTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2865	GAGCAACCTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	2924
Qy	2702	TAAATCCATGTCTGGAGCGCTCAAGAGTGTGGTGTGTCTCCAGCTGTAGATGAAA	2761
Db	2925	TAAATCCATGTCTGGAGCGCTCAAGAGTGTGGTGTGTCTCCAGCTGTAGATGAAA	2984
Qy	2762	ATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATTTAGACAGACC	2821
Db	2985	ATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATTTAGACAGACC	3044
Qy	2822	ATTATGAGTCTCAAGGAGAGGAGATACCTTACCTTAAAGTCAAGCCAGAAAAAGTGGTT	2881
Db	3045	ATTATGAGTCTCAAGGAGAGGAGATACCTTACCTTAAAGTCAAGCCAGAAAAAGTGGTT	3104
Qy	2882	TCCAAATGGAATGCTGTCTGAACTCCACAGTGAAGCCACAGCTTTTATTTGGGACCCAGG	2941
Db	3105	TCCAAATGGAATGCTGTCTGAACTCCACAGTGAAGCCACAGCTTTTATTTGGGACCCAGG	3164
Qy	2942	TCCTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAAGCCTTTCTTTG	3001
Db	3165	TCCTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAAGCCTTTCTTTG	3224
Qy	3002	ATGTCTGGAGCTCCGGGGCAAGTACCCGAATGAGAGCTTCCCAAGATATTTAACGACA	3061
Db	3225	ATGTCTGGAGCTCCGGGGCAAGTACCCGAATGAGAGCTTCCCAAGATATTTAACGACA	3284
Qy	3062	AAACAGTAGTGGAGGCGCAGGAGGCTCTACGATGATGCCCAATATGCTGAACACAC	3121
Db	3285	AAACAGTAGTGGAGGCGCAGGAGGCTCTACGATGATGCCCAATATGCTGAACACAC	3344
Qy	3122	TGATTAGTCAAAAGAAACTCCGGGCCCGGGTGTGGTGTGGTCTTGCCCAAGCAAGATA	3181
Db	3345	TGATTAGTCAAAAGAAACTCCGGGCCCGGGTGTGGTGTGGTCTTGCCCAAGCAAGATA	3404
Qy	3182	TCCAAGACGACATTCACCTGTACGAGAGCTGTGTGTCGCCAGGCTGCGAGGCCATAG	3241
Db	3405	TCCAAGACGACATTCACCTGTACGAGAGCTGTGTGTCGCCAGGCTGCGAGGCCATAG	3464
Qy	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACCGGAGCCATCT	3301
Db	3465	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACCGGAGCCATCT	3524

```
Qy 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTCAGTACCTGGCCCTGT 3361
Db |
Qy 3525 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTCAGTACCTGGCCCTGT 3584
Db |
Qy 3362 TTGCCGTTGCTGCTTTGGGTAGAGAGCTGAGCAAGCCCTATGAGGATGATGTTGACG 3421
Db |
Qy 3585 TTGCCGTTGCTGCTTTGGGTAGAGAGCTGAGCAAGCCCTATGAGGATGATGTTGACG 3644
Db |
Qy 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGCTGGCAGAGCCCTTTCAGAG 3481
Db |
Qy 3645 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGCTGGCAGAGCCCTTTCAGAG 3704
Db |
Qy 3482 AGCTCCATGAAGAGTTCCCGCAGAACTGTGGGCTACTGTGGCAGTGCAGAGCTGTCAGC 3541
Db |
Qy 3705 AGCTCCATGAAGAGTTCCCGCAGAACTGTGGGCTACTGTGGCAGTGCAGAGCTGTCAGC 3764
Db |
Qy 3542 TCGCAGACTCGCAAGGTTGGGTCACAAAGGATCCCGCCCGCTCTGGCTACCCAGCC 3601
Db |
Qy 3765 TCGCAGACTCGCAAGGTTGGGTCACAAAGGATCCCGCCCGCTCTGGCTACCCAGCC 3824
Db |
Qy 3602 AGCCGACACACCGAGAGCTCCACATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
Db |
Qy 3825 AGCCGACACACCGAGAGCTCCACATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3884
Db |
Qy 3662 GCATTAGGTTAACAGAACTATTAGCAATGGCACTGTCTTCAGCAGTCTCAGGCTCTACT 3721
Db |
Qy 3885 GCATTAGGTTAACAGAACTATTAGCAATGGCACTGTCTTCAGCAGTCTCAGGCTCTACT 3944
Db |
Qy 3722 TCTCAATTTGAAGTCCAAATATTGTTGCTGGGGAAGATTTCGAAGATCAGGTTGAGG 3781
Db |
Qy 3945 TCTCAATTTGAAGTCCAAATATTGTTGCTGGGGAAGATTTCGAAGATCAGGTTGAGG 4004
Db |
Qy 3782 ATTATGCAATGAGGAGAACATATCTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841
Db |
Qy 4005 ATTATGCAATGAGGAGAACATATCTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4064
Db |
Qy 3842 TGGGATATGATACAGACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3901
Db |
Qy 4065 TGGGATATGATACAGACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4122
Db |
Qy 3902 CAAGGAATATACAACTAG 3919
Db 4123 CAAGGAATATACAACTAG 4140
```

```
RESULT 7
HSU75743 7122 bp mRNA linear PRI 06-JUN-1997
LOCUS Human methionine synthase mRNA, complete cds.
DEFINITION
ACCESSION U75743
VERSION U75743.1 GI:1763339
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 7122)
AUTHORS Li,Y.N., Gulati,S., Baker,P.J., Brody,L.C., Banerjee,R. and
Kruger,W.D.
TITLE Cloning, mapping, and RNA analysis of the human methionine synthase
gene
JOURNAL Hum. Mol. Genet. (1996) In press
REFERENCE
2 (bases 1 to 7122)
AUTHORS Li,Y.N., Gulati,S., Baker,P.J., Brody,L.C., Banerjee,R. and
Kruger,W.D.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1996) Population Science, Fox Chase Cancer
Center, 7701 Burholme Ave., Philadelphia, PA 19111, USA
FEATURES
1..7122
source /organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db xref="taxon:9606"
/chromosome="1"
/map="1q42-44"
/cell_type="lymphoid"
/notes="composite cdna"
287..4084
/notes="similar to Escherichia coli methionine synthase
encoded by GenBank Accession Number J04975"
/codon_start=1
/product="methionine synthase"
/protein_id="AAB58906.1"
/db xref="GI:216099"
/translation="MSPALQDSQPEGLKTKLRLDEINAIQKRMVLVGGMTMIQRE
KLAHEPRGGEFKDHPKPLKGNNDILSIQTPDVIYQIHKEYLLAGADIETNTSSST
IAQADVLEHLAVKMMSCSAGVAKAEVTLQGIKRFVAGALGPTKNTLSVSPSVS
RPRYRNTPELVEAQEQAKGLLDGVDILLIETPTDTANAKAALFALQNLPEKVA
PRPFIISGTTIVDSKGRSLSGQTEGPFVISHGEPYLYIGLNCALGAAPFPFIRIK
CTTAYLYPNAGLPNTGDDYDETSMKAKHLDKFAMDGLVNIIVGGCCGTPDHIREI
AEAVKNCPRVPATAPEGHMLLSGLEPFPRIGPYTNFVNIGERCNVAAGSRKFAKLINA
GNYEALCVAKVOVMGAQVLDVNMDDGLDGPSTMTFCNLIASEPDIKAPLCLIDS
SNFAVLEAGLKCQGGKLVNSISLKSGEDDPLEKARKIKYGAAMVVMVAFDEBQATE
TDRKIRVCTRAYHLLVKLGFNPNDIIFDPNIIITIGTMEHNLVAINFIHATKVIKE
TLPGARISGGLSNLSFSFRGEAIRAMHGVFLYHAIKSGMDMGINVAGNLVYDDLIH
KELLQCEDLIWNKDPEATEKLLRYAQQTGGKKVQTDDEWRNGPVEELEYALVVG
IEKHIIEDTEARLNQKYPRLNIIEGFLMNGMKIVGDLFGAGKMLPQVITKSARVM
KKAVGHLIIPMEKERBETRLNGTVEEDPYQGTIVLATVKGVDHDIKGNIVGVLCG
NNFRVLDLGMWTPCDKILKAALDHKADIIGLSGLITPSLDEMILFVAKEMERLARIPL
LIGATTSKTHAVTAIRYSAPVHIVLDASKSVVCSQLLDENLKDYEEMEVEEYB
DIRQDHVESLKKERRYLPISQARKSGFQMDLSBPFPVKPTFTPTQVPEVDLQKLDY
IDGKFPFVQMLRKYPNRGFPKIFNDKTGGVARKVYDDAHNMLNTLSIQKLRARG
LVGFWPAQSIQDDIHLYAEAAVPAQAEPIATFYGLRQQAEDKSDASTPEYTCLSDFIAP
WHSIRDLGLFAVACPVSEELSKAVEDDDGYSSIMVKGALGDLRAFAELHVERV
REUMVACGSQLDVLADRLRLRYKIRPAPGYPSPQPHTEKLTWMLADIIQSQTGIRLT
ESLAMAPASAVSGLIFSNLSKIFAVGKISKOQVEDYALRNKNSVAEVWKUGPILGY
DTD"
```

Query Match		99.18	Score 3884.8	DB 8	Length 7122
Best Local Similarity		99.6	Pred. No. 0		
Matches 3904		Conservative 0	Mismatches 12	Indels 2	Gaps 1
Qy	2	GTCACTGTGTGAGAGACGCTCTCTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	61		
Db	225	GTCACTGTGTGAGAGACGCTCTCTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	284		
Qy	62	ACATGTACCCGCGCTCCAGACCTGTGCGAACCCGAGAGTCTGCAAGAAACCTCTGGGG	121		
Db	285	ACATGTACCCGCGCTCCAGACCTGTGCGAACCCGAGAGTCTGCAAGAAACCTCTGGGG	344		
Qy	122	ATGAGATCAATGTCATCTTCGAGAAAGAGGATTATGGTGTGGATGGAGGGATGGGACCA	181		
Db	345	ATGAGATCAATGTCATCTTCGAGAAAGAGGATTATGGTGTGGATGGAGGGATGGGACCA	404		
Qy	182	TGATCCAGCGGGAGAGCTAAACGAGAACTTTCCGAGGTCAGGAAATTTAAAGATCATG	241		
Db	405	TGATCCAGCGGGAGAGCTAAACGAGAACTTTCCGAGGTCAGGAAATTTAAAGATCATG	464		
Qy	242	CCAGGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCATTACC	301		
Db	465	CCAGGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCATTACC	524		
Qy	302	AAATCCATAGGAATACCTTGCTGGCTGGGCGAGATATCATTTGAAACAAACTTTTAGCA	361		
Db	525	AAATCCATAGGAATACCTTGCTGGCTGGGCGAGATATCATTTGAAACAAACTTTTAGCA	584		
Qy	362	GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAACACTTTGGCCTACCGGATGAACATGT	421		
Db	585	GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAACACTTTGGCCTACCGGATGAACATGT	644		
Qy	422	GCTCTGACAGGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	481		
Db	645	GCTCTGACAGGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	704		

ORIGIN

482 GGTGTCGAGGGCTCTGGTCCGACTAATAGACACTCTCTGTGTCCCATCTGTGG 541  
Db  
705 GGTGTCGAGGGCTCTGGTCCGACTAATAGACACTCTCTGTGTCCCATCTGTGG 764  
Qy  
542 AAAGCCGGATTAAGGAAATCAATCATCTGATGAGCTTTGTAAGCATACCAAGAGCAG 601  
Db  
765 AAAGCCGGATTAAGGAAATCAATCATCTGATGAGCTTTGTAAGCATACCAAGAGCAG 824  
Qy  
602 CCAAGGACTCTTGATGAGGGGTTGATATCTTATCTCATCTGAACTATTTTGTGATCTG 661  
Db  
825 CCAAGGACTCTTGATGAGGGGTTGATATCTTATCTCATCTGAACTATTTTGTGATCTG 884  
Qy  
662 CCAATGCCAAGGAGCCTTTGTTGCACTCCAAATCTTTTGGAGGAATATGCTCCCC 721  
Db  
885 CCAATGCCAAGGAGCCTTTGTTGCACTCCAAATCTTTTGGAGGAATATGCTCCCC 944  
Qy  
722 GGCCTATCTTTATTTCAAGGAGCAGCTGTTGATATAAAGTGGCGGACTCTTTCCGGACAGA 781  
Db  
945 GGCCTATCTTTATTTCAAGGAGCAGCTGTTGATATAAAGTGGCGGACTCTTTCCGGACAGA 1004  
Qy  
782 CAGGAGGAGGATTTGTCATCAGCGTGTCTCATGAGGAACCACTCTGCAATTTGGAATTA 841  
Db  
1005 CAGGAGGAGGATTTGTCATCAGCGTGTCTCATGAGGAACCACTCTGCAATTTGGAATTA 1064  
Qy  
842 GTGCTTTGGTGCAGCTGAGATGAGACTTTTATTAATTAATTTGGAATTTGGAATTTGGA 901  
Db  
1065 GTGCTTTGGTGCAGCTGAGATGAGACTTTTATTAATTAATTTGGAATTTGGAATTTGGA 1124  
Qy  
902 CCTATGCTCTGTTATCCCAATGCAAGTCTTTCCCAACACTTTGGTGACTATGATGAAA 961  
Db  
1125 CCTATGCTCTGTTATCCCAATGCAAGTCTTTCCCAACACTTTGGTGACTATGATGAAA 1184  
Qy  
962 CGCCTTCTATGATGGCCCAAGCACCATAAGGATTTTGTATGATGGCTTGGTCAATATAG 1021  
Db  
1185 CGCCTTCTATGATGGCCCAAGCACCATAAGGATTTTGTATGATGGCTTGGTCAATATAG 1244  
Qy  
1022 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db  
1245 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1304  
Qy  
1082 ATTGTAAGCTTAGAGTTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGTC 1141  
Db  
1305 ATTGTAAGCTTAGAGTTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGTC 1364  
Qy  
1142 TAGAGCCCTTCAGGATTTGGACCGTACACCAACTTTTGTAACTTTGGAGAGCGCTGTAATG 1201  
Db  
1365 TAGAGCCCTTCAGGATTTGGACCGTACACCAACTTTTGTAACTTTGGAGAGCGCTGTAATG 1424  
Qy  
1202 TTGAGGATCAAGGAAGTTTGTCTAACTCATCATGCGAGGAACATATGAAGAGCCTTGT 1261  
Db  
1425 TTGAGGATCAAGGAAGTTTGTCTAACTCATCATGCGAGGAACATATGAAGAGCCTTGT 1484  
Qy  
1262 GTGTTGCCAAGTGCAGGTGGAATGGAGCCAGCTGTTGGATGTCACATGATGATGATG 1321  
Db  
1485 GTGTTGCCAAGTGCAGGTGGAATGGAGCCAGCTGTTGGATGTCACATGATGATGATG 1544  
Qy  
1322 GCATGCTAGATGTTCCAAAGTGCATGACCAAGATTTTGCACCTTAATTTGCTTCCAGGCCAG 1381  
Db  
1545 GCATGCTAGATGTTCCAAAGTGCATGACCAAGATTTTGCACCTTAATTTGCTTCCAGGCCAG 1604  
Qy  
1382 ACATGCGAAGGATACCTTTTGTGCACTGACCTCCCAATTTTGTGCTGATGAACTGGGT 1441  
Db  
1605 ACATGCGAAGGATACCTTTTGTGCACTGACCTCCCAATTTTGTGCTGATGAACTGGGT 1664  
Qy  
1442 TAAAGTGTGCCAAGGAGTGCATTTGTCATACATTTAGTCTGAAGGAGGAGGAGCAG 1501  
Db  
1665 TAAAGTGTGCCAAGGAGTGCATTTGTCATACATTTAGTCTGAAGGAGGAGGAGCAG 1724  
Qy  
1502 ACTTCTTGAGAGGCGCAGGAAGATTAAGAAAGTATGAGCTGCTATGCTGTGATGCTT 1561  
Db  
1725 ACTTCTTGAGAGGCGCAGGAAGATTAAGAAAGTATGAGCTGCTATGCTGTGATGCTT 1784  
Qy  
1562 TTGATGAAGAGGAGCAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCACCCGGCCT 1621

1785 TTGATGAAGAGGAGCAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCACCCGGCCT 1844  
Qy  
1622 ACCATCTGCTGTGAAAAAATCTGGCTTTAATCAAATGACATTAATTTTGGACCTTAATA 1681  
Db  
1845 ACCATCTGCTGTGAAAAAATCTGGCTTTAATCAAATGACATTAATTTTGGACCTTAATA 1904  
Qy  
1682 TCTTAACCATTTGGGACTGGAAATGGAGGAACAACACTTTGTATGCCATTAATTTTATCCATG 1741  
Db  
1905 TCTTAACCATTTGGGACTGGAAATGGAGGAACAACACTTTGTATGCCATTAATTTTATCCATG 1964  
Qy  
1742 CAAACAAAGTCATTAAGAAAAATTAACCTGGAGCCAGAAATAAGTGGAGTCTTTTCCAACT 1801  
Db  
1965 CAAACAAAGTCATTAAGAAAAATTAACCTGGAGCCAGAAATAAGTGGAGTCTTTTCCAACT 2024  
Qy  
1802 TGTCTTCTCTCTCCGAGGAATGGAAGCCATTTCCGAGGAACAATGCAATGGGGTTTTCTTT 1861  
Db  
2025 TGTCTTCTCTCTCCGAGGAATGGAAGCCATTTCCGAGGAACAATGCAATGGGGTTTTCTTT 2084  
Qy  
1862 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTTCTGTGT 1921  
Db  
2085 ACCATGCAATCAAGTCTGGCATGGAGATAGTGAATGCTGGAAACCTTCTGTGT 2144  
Qy  
1922 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 1981  
Db  
2145 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 2204  
Qy  
1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2041  
Db  
2205 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2264  
Qy  
2042 TCATTGAGACTGATGAGTGGAGAAATGGCCCTGTCGAGAAACGCTTGTAGTATGCCCTTG 2101  
Db  
2265 TCATTGAGACTGATGAGTGGAGAAATGGCCCTGTCGAGAAACGCTTGTAGTATGCCCTTG 2224  
Qy  
2102 TGAAGGGCATTTGAAGAACATATTTTGAAGTACTGAGGAGCCAGCTTAAACCAAAAA 2161  
Db  
2325 TGAAGGGCATTTGAAGAACATATTTTGAAGTACTGAGGAGCCAGCTTAAACCAAAAA 2384  
Qy  
2162 AATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
Db  
2385 AATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2444  
Qy  
2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTATAAAGTCAAGCCCGGTTA 2281  
Db  
2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTATAAAGTCAAGCCCGGTTA 2504  
Qy  
2282 TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACCCAGAG 2341  
Db  
2505 TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACCCAGAG 2364  
Qy  
2342 TGCTTAAACGGCACAAGTAGAAGAGAGAGCCCTTACAGGGGACCATCTGCTGGCCACTG 2401  
Db  
2565 TGCTTAAACGGCACAAGTAGAAGAGAGAGCCCTTACAGGGGACCATCTGCTGGCCACTG 2624  
Qy  
2402 TTAAGCGCAGCTGTCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2461  
Db  
2625 TTAAGCGCAGCTGTCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2684  
Qy  
2462 ATTTCCGAGTTATTTGATTTAGGAGTCAATGATCCATGATGATAGATGATGAAAGCTGCTC 2521  
Db  
2685 ATTTCCGAGTTATTTGATTTAGGAGTCAATGATCCATGATGATAGATGATGAAAGCTGCTC 2744  
Qy  
2522 TTGACCAACAAAGCAGATATTAATTTGGCTGTGAGACTCATCTCTTCCCTGGATGAAA 2581  
Db  
2745 TTGACCAACAAAGCAGATATTAATTTGGCTGTGAGACTCATCTCTTCCCTGGATGAAA 2804  
Qy  
2582 TGAATTTTGTGCAAGGAATTTAGGAGATTTAGCTATAAGGATTTCCATTTGATTTGGAG 2641  
Db  
2805 TGAATTTTGTGCAAGGAATTTAGGAGATTTAGCTATAAGGATTTCCATTTGATTTGGAG 2864  
Qy  
2642 GAGCAACCACTTCAAAAAACCCACAGAGCAGTTAAAAATAGCTCCGAGTATACAGTGCACCTG 2701



Db 2865 GAGCAACCACTTCAAAAAACCCACACAGCAGTAAATAATAGCTCCGAGATACAGTGCACCTG 2924  
Qy 2702 TAAATCCATGCTCTGACGCGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761  
Db 2925 TAAATCCATGCTCTGACGCGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2984  
Qy 2762 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAATATCAAGATATTAGACAGGACC 2821  
Db 2985 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAATATGAAGATATTAGACAGGACC 3044  
Qy 2822 ATTATGAGTCTCTAAGAGAGGAGATACCTTAACCTTAAGTCAAGCCAGAAAAAGTGTGT 2881  
Db 3045 ATTATGAGTCTCTAAGAGAGGAGATACCTTAACCTTAAGTCAAGCCAGAAAAAGTGTGT 3104  
Qy 2882 TCCAAATGGATTGGTGTCTGAACCTCACCAAGTGAAGCCACGCTTTATTGCGACCCAGG 2941  
Db 3105 TCCAAATGGATTGGTGTCTGAACCTCACCAAGTGAAGCCACGCTTTATTGCGACCCAGG 3164  
Qy 2942 TCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAAAGCTTTCTTTG 3001  
Db 3165 TCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAAAGCTTTCTTTG 3224  
Qy 3002 ATGCTCTGGAGCTCCGGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA 3061  
Db 3225 ATGCTCTGGAGCTCCGGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA 3284  
Qy 3062 AAACAGTAGGTGGAGAGGCCAGGAGGCTACGATGATGCCACAAATATCTGTAACACAC 3121  
Db 3285 AAACAGTAGGTGGAGAGGCCAGGAGGCTACGATGATGCCACAAATATCTGTAACACAC 3344  
Qy 3122 TGATTAGTCAAAAGAAATCCCGGGCCCGGGTGTGGTTGGGTTCTGGCCAGCACAGAGTA 3181  
Db 3345 TGATTAGTCAAAAGAAATCCCGGGCCCGGGTGTGGTTGGGTTCTGGCCAGCACAGAGTA 3404  
Qy 3182 TCCAAAGCAGCATTCCTGTATCGCAGAGGCTGTGTGCCAGGCTGCAGAGGCCATPAG 3241  
Db 3405 TCCAAAGCAGCATTCCTGTATCGCAGAGGCTGTGTGCCAGGCTGCAGAGGCCATPAG 3464  
Qy 3242 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGCCATACT 3301  
Db 3465 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGCCATACT 3524  
Qy 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACTTGGGCTGT 3361  
Db 3525 ACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACTTGGGCTGT 3584  
Qy 3362 TTGCCGTTGCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGCCCTATGAGGATGATGTGACG 3421  
Db 3585 TTGCCGTTGCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGCCCTATGAGGATGATGTGACG 3644  
Qy 3422 ACTACAGCAGCATCATGTTCAAGGCGCTGGGGGACCGGCTGCCAGAGGCTTTGCAGAG 3481  
Db 3645 ACTACAGCAGCATCATGTTCAAGGCGCTGGGGGACCGGCTGCCAGAGGCTTTGCAGAG 3704  
Qy 3482 AGCTCCATGAAGAGTTGCCCGAGAACTGTGGGCTACTGTGGCAGTGTGAGCAGCTGGACG 3541  
Db 3705 AGCTCCATGAAGAGTTGCCCGAGAACTGTGGGCTACTGTGGCAGTGTGAGCAGCTGGACG 3764  
Qy 3542 TCGCAGACTCGCAAGGTTGCGGTACAAGGGCATCCGCCGCTCTTGGCTACCCCGCC 3601  
Db 3765 TCGCAGACTCGCGAGGCTGCGGTACAAGGGCATCCGCCGCTCTTGGCTACCCCGCC 3824  
Qy 3602 AGCCGACCCACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGAGTCTACAG 3661  
Db 3825 AGCCGACCCACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGAGTCTACAG 3884  
Qy 3662 GCATTAGGTTAAACAGAACTATTAGCAATGGCACTCTTACAGAGTCTCAGGCTCTACT 3721  
Db 3885 GCATTAGGTTAAACAGAACTATTAGCAATGGCACTCTTACAGAGTCTCAGGCTCTACT 3944  
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGATCAAGGTTGAGG 3781  
Db 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGATCAAGGTTGAGG 4004

Qy 3782 ATTATGCAATTGAGGAACAATATCTGTGCTGAGGTTGAGAAATGCTTGGACCCATTT 3841  
Db 4005 ATTATGCAATTGAGGAACAATATCTGTGCTGAGGTTGAGAAATGCTTGGACCCATTT 4064  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGGCTTTTATCTTTGATGATCCT 3901  
Db 4065 TGGGATATGATACAGACTAAC--TTTTTTTTTTTTTTTGGCTTTTATCTTTGATGATCCT 4122  
Qy 3902 CAAGGAATACAACTAG 3919  
Db 4123 CAAGGAATACAACTAG 4140  
RESULT 8  
AR367908  
LOCUS AR367908 7224 bp DNA linear PAT 12-SEP-2003  
DEFINITION Sequence 6 from patent US 6376210.  
ACCESSION AR367908  
VERSION AR367908.1 GI:34601379  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7224)  
AUTHORS Yuan,C.-S.  
TITLE Methods and compositions for assaying analytes  
JOURNAL Patent: US 6376210-A 6 23-APR-2002;  
General Atomics; San Diego, CA  
FEATURES  
source 1..7224  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.1%; Score 3883.2; DB 6; Length 7224;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3903; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
Qy 2 GTACCTGTGGAGAGCAGCTTCTCTGCCGCCCTCTGGCAAGGAGGAGCTCGACA 61  
Db 333 GTACCTGTGGAGAGCAGCTTCTCTGCCGCCCTCTGGCAAGGAGGAGCTCGACA 392  
Qy 62 ACATGTCAACCGCTCCAGACCTGTGCAACCCGAGGCTCGAAGAAACCTCGGGG 121  
Db 393 ACATGTCAACCGCTCCAGACCTGTGCAACCCGAGGCTCGAAGAAACCTCGGGG 452  
Qy 122 ATGAGATCAATGCCAATTTCTGCAGAAAGGATTTATGGTCTGGATGGAGGATGGGACCA 181  
Db 453 ATGAGATCAATGCCAATTTCTGCAGAAAGGATTTATGGTCTGGATGGAGGATGGGACCA 512  
Qy 182 TGATCCAGCGGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241  
Db 513 TGATCCAGCGGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 572  
Qy 242 CCAGCGCTCAAAAGGCAACAATGACATTTTAAAGTATAAATCAGCTGATGCTATTACC 301  
Db 573 CCAGCGCTCAAAAGGCAACAATGACATTTTAAAGTATAAATCAGCTGATGCTATTACC 632  
Qy 302 AAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 361  
Db 633 AAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 692  
Qy 362 GCATAGTATTGCCAGCTGACTATGGCTTTGAACACTTGGCTTACCGGATGACATGT 421  
Db 693 GCATAGTATTGCCAGCTGACTATGGCTTTGAACACTTGGCTTACCGGATGACATGT 752  
Qy 422 GCTCTCAGGAGTGGCCAGAAAGCTGCCAGGAGGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTCAGGAGTGGCCAGAAAGCTGCCAGGAGGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGTGGCAGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG 541

Db 813 GGTGTTGGCAGGGGCTCTGGGTCGACTAATAGACACTCTCTGTGTCCCATCTGTGG 872  
Qy 542 AAAGCCGGATTATAGGAACATCACTTGTATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
Db 873 AAAGCCGGATTATAGGAACATCACTTGTATGAGCTTGTGAAGCATACCAAGAGCAGG 932  
Qy 602 CCAAAGGACTTCTGGATGGGGGTGTGATATCTTACTCATTTGAACATATTTTGTATGACTG 661  
Db 933 CCAAAGGACTTCTGGATGGGGGTGTGATATCTTACTCATTTGAACATATTTTGTATGACTG 992  
Qy 662 CCAATGCCAAGCAGCCTTGTTCACCTCCAAATCTTTTGTAGGAGAAATATGCTCCCC 721  
Db 993 CCAATGCCAAGCAGCCTTGTTCACCTCCAAATCTTTTGTAGGAGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTTTCAGGAGCATGTGTGATAAAAGTGGCGGACCTCTTTCGGACAGA 781  
Db 1053 GGCCTATCTTTATTTTCAGGAGCATGTGTGATAAAAGTGGCGGACCTCTTTCGGACAGA 1112  
Qy 782 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTGCTATGGATTAAAT 841  
Db 1113 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTGCTATGGATTAAAT 1172  
Qy 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATAAATTTGGAATAATGTACAACAG 901  
Db 1173 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATAAATTTGGAATAATGTACAACAG 1232  
Qy 902 CCTATGCTCTCTGTTATCCCAATGCAGGTCTTCCAAACACCTTTGGTGTGACTATGATGAAA 961  
Db 1233 CCTATGCTCTCTGTTATCCCAATGCAGGTCTTCCAAACACCTTTGGTGTGACTATGATGAAA 1292  
Qy 962 CGCCTCTATGATGGCCCAAGCAGCTTAAGGATTTTGTATGAGTGGCTTGGTCAATATAG 1021  
Db 1293 CGCCTCTATGATGGCCCAAGCAGCTTAAGGATTTTGTATGAGTGGCTTGGTCAATATAG 1352  
Qy 1022 TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTGTAAGCTTAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141  
Db 1413 ATTGTAAGCTTAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1472  
Qy 1142 TAGAGCCCTTCAGATTTGGACCGTACACCAACTTTTGTATGAGTGGAGCGCTGTAAATG 1201  
Db 1473 TAGAGCCCTTCAGATTTGGACCGTACACCAACTTTTGTATGAGTGGAGCGCTGTAAATG 1532  
Qy 1202 TTGCAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACATATGAAGAGCCCTTGT 1261  
Db 1533 TTGCAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACATATGAAGAGCCCTTGT 1592  
Qy 1262 GTGTGCCAAAGTGCAGGTGGAATGGGAGCCCAAGGTGTTGGATGTCAACATGAGATGATG 1321  
Db 1593 GTGTGCCAAAGTGCAGGTGGAATGGGAGCCCAAGGTGTTGGATGTCAACATGAGATGATG 1652  
Qy 1322 GCATGCTAGATGTCCAAAGTGCATATGACAGATTTTGCACCTTAAATTTGCTTCGAGCCAG 1381  
Db 1653 GCATGCTAGATGTCCAAAGTGCATATGACAGATTTTGCACCTTAAATTTGCTTCGAGCCAG 1712  
Qy 1382 ACATCGCAAGGTACTTTTGTGCAATCGACTCTCTCCAAATTTTGTGCTGATTTGAGCTGGGT 1441  
Db 1713 ACATCGCAAGGTACTTTTGTGCAATCGACTCTCTCCAAATTTTGTGCTGATTTGAGCTGGGT 1772  
Qy 1442 TAAAGTGTGTCGAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCG 1501  
Db 1773 TAAAGTGTGTCGAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCG 1832  
Qy 1502 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGGCTGTATGGTGGTCAATGGCTT 1561  
Db 1833 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGGCTGTATGGTGGTCAATGGCTT 1892  
Qy 1562 TTGATGAGAGGACAGGCAACAGAACAGACCAAAATCAGAGTGTGCACCGGGCCT 1621  
Db 1893 TTGATGAGAGGACAGGCAACAGAACAGACCAAAATCAGAGTGTGCACCGGGCCT 1952

Qy 1622 ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATATTTTGTACCCCTAATA 1681  
Db 1953 ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATATTTTGTACCCCTAATA 2012  
Qy 1682 TCCTAACCATTGGGACTGGAAATGGAGGAACAACAATCTTGTATGCAATTAATTTTATCCATG 1741  
Db 2013 TCCTAACCATTGGGACTGGAAATGGAGGAACAACAATCTTGTATGCAATTAATTTTATCCATG 2072  
Qy 1742 CAACAAAGTGCATTTAAAGAAACATTAACCTGGAGCCAGAAATGCAATGGGGTTTCCCTTT 1801  
Db 2073 CAACAAAGTGCATTTAAAGAAACATTAACCTGGAGCCAGAAATGCAATGGGGTTTCCCTTT 2132  
Qy 1802 TGTCTCTCTCTTCCGAGGAAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTCCCTTT 1861  
Db 2133 TGTCTCTCTCTTCCGAGGAAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTCCCTTT 2192  
Qy 1862 ACCATGCAATCAAGTCTGGCATGCAATGAGATAGTGAATGCTGGAACCTCCCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGGCATGCAATGAGATAGTGAATGCTGGAACCTCCCTGTGT 2252  
Qy 1922 ATGATGATATCCATAAGGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAAC 1981  
Db 2253 ATGATGATATCCATAAGGAACTTCTGAGCTCTGTGAGATCTCATCTGGAATAAGAAC 2312  
Qy 1982 CTGAGGCCACTGAGAAAGCTTTAGCTTATGCCCAGACTCAAGGCAACAGAGGGAAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAAAGCTTTAGCTTATGCCCAGACTCAAGGCAACAGAGGGAAGAAAG 2372  
Qy 2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAACCCCTTGAATGATGCCCCCTG 2101  
Db 2373 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAACCCCTTGAATGATGCCCCCTG 2432  
Qy 2102 TGAAGGGCATTTGAAAAACATATTTATGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA 2161  
Db 2433 TGAAGGGCATTTGAAAAACATATTTATGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA 2492  
Qy 2162 AATATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
Db 2493 AATATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTTAAAGTCAAGCCGGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTTAAAGTCAAGCCGGGTTA 2612  
Qy 2282 TGAAGAGGCTGTGTGGCCACCTTATCCCTTTTCAATGGAAGAAAGAGAGAAACCAAGAG 2341  
Db 2613 TGAAGAGGCTGTGTGGCCACCTTATCCCTTTTCAATGGAAGAAAGAGAGAAACCAAGAG 2672  
Qy 2342 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACAGGGCACCATGCTGCTGGCCACTG 2401  
Db 2673 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACAGGGCACCATGCTGCTGGCCACTG 2732  
Qy 2402 TTAAGGGCAGCTGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCATATA 2461  
Db 2733 TTAAGGGCAGCTGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCATATA 2792  
Qy 2462 ATTTCCGAGTTATTTGATTTAGGAGTCAATGACTCCATGTAAGATACATGAAAGCTGCTC 2521  
Db 2793 ATTTCCGAGTTATTTGATTTAGGAGTCAATGACTCCATGTAAGATACATGAAAGCTGCTC 2852  
Qy 2522 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTCTCCCTGGATGAAA 2581  
Db 2853 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTCTCCCTGGATGAAA 2912  
Qy 2582 TGATTTTGTGTCGAAGGAAATGGAGAGATTAGCTATAGGATTCATTTGTTGATTTGGAG 2641  
Db 2913 TGATTTTGTGTCGAAGGAAATGGAGAGATTAGCTATAGGATTCATTTGTTGATTTGGAG 2972  
Qy 2642 GAGCAACCACTTCCAAAAACCCACACAGCAGTGTAAAAATAGCTCCGAGATACAGTGACCTG 2701  
Db 2973 GAGCAACCACTTCCAAAAACCCACACAGCAGTGTAAAAATAGCTCCGAGATACAGTGACCTG 3032

QY 2702 TAATCCATGTCCTCGACGCGTCCAAGAGTGTGGTGGTGTGTCCAGCTGTTAGATGAAA 2761  
Db 3033 TAATCCATGTCCTCGACGCGTCCAAGAGTGTGGTGGTGTGTCCAGCTGTTAGATGAAA 3092  
QY 2762 ATCTAAAGGATGAATACCTTGAAGGAATCATGGAAGATATGAAGATATTAAGACGACC 2821  
Db 3093 ATCTAAAGGATGAATACCTTGAAGGAATCATGGAAGATATGAAGATATTAAGACGACC 3152  
QY 2822 ATTATGAGTCTCTCAAGAGAGGAGATACCTTAAAGTCAAGCCAGAAAGTGGTT 2881  
Db 3153 ATTATGAGTCTCTCAAGAGAGGAGATACCTTAAAGTCAAGCCAGAAAGTGGTT 3212  
QY 2882 TCCAAATGGATTGGTGTCTGAACCTCACCCAGTGAAGCCACGCTTTATTGGGACCCAGG 2941  
Db 3213 TCCAAATGGATTGGTGTCTGAACCTCACCCAGTGAAGCCACGCTTTATTGGGACCCAGG 3272  
QY 2942 TCTTTGAAGATATGACTCGAGAAAGCTGGTGGACTCATATTGACTGGAAGCCTTCTTTTG 3001  
Db 3273 TCTTTGAAGATATGACTCGAGAAAGCTGGTGGACTCATATTGACTGGAAGCCTTCTTTTG 3332  
QY 3002 ATGCTCGACGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAACGACA 3061  
Db 3333 ATGCTCGACGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAACGACA 3392  
QY 3062 AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCCAACAATATGCTGAACACAC 3121  
Db 3393 AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCCAACAATATGCTGAACACAC 3452  
QY 3122 TGATTAGTCAAAAGAACTCCGGGCGCGGGGTGGTGGTCTGGCCAGCAGCAGGATA 3181  
Db 3453 TGATTAGTCAAAAGAACTCCGGGCGCGGGGTGGTGGTCTGGCCAGCAGCAGGATA 3512  
QY 3182 TCCAGACGACATTCACCTGTACGAGAGGCTGCTGCGCCAGGCTCGACAGCCCATAG 3241  
Db 3513 TCCAGACGACATTCACCTGTACGAGAGGCTGCTGCGCCAGGCTCGACAGCCCATAG 3572  
QY 3242 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCAGCCATACT 3301  
Db 3573 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCAGCCATACT 3632  
QY 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTGGCATCCGTGACTACCTGGGCCCTGT 3361  
Db 3633 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTGGCATCCGTGACTACCTGGGCCCTGT 3692  
QY 3362 TTGCGGTTGCTGCTTGGGGTAGAAGAGCTGAGCAAGCCCTTACGAGATGATGGTGACG 3421  
Db 3693 TTGCGGTTGCTGCTTGGGGTAGAAGAGCTGAGCAAGCCCTTACGAGATGATGGTGACG 3752  
QY 3422 ACTACAGCAGCATCATGCTCAAGCGGCTGGGGGACCGGCTGGCAGAGCGCTTTGCAGAG 3481  
Db 3753 ACTACAGCAGCATCATGCTCAAGCGGCTGGGGGACCGGCTGGCAGAGCGCTTTGCAGAG 3812  
QY 3482 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGCTGGAGC 3541  
Db 3813 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGCTGGAGC 3872  
QY 3542 TCGCAGACTCGAAGGTTGCGGTACAGGCAATCCCGCGGCTCCTGGCTACCCAGCC 3601  
Db 3932 TCGCAGACTCGCAGGCTGCGGTACAGGCAATCCCGCGGCTCCTGGCTACCCAGCC 3932  
QY 3602 AGCCCGACACACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
Db 3993 AGCCCGACACACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3992  
QY 3662 GCATTAGGTTAAGCAATCATTAGCAATGGCACTGCTTCAAGCAGTCTCAGCCCTTACT 3721  
Db 3993 GCATTAGGTTAAGCAATCATTAGCAATGGCACTGCTTCAAGCAGTCTCAGCCCTTACT 4052  
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGAAGTCCAAATATTTTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 4112  
QY 3782 ATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGGCTTGGACCCATTT 3841

Db 4113 ATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTGGACCCATTT 4172  
QY 3842 TGGGATATGATACAGACTAACTTTTGTGCTGCTTTTGTGCTTTTATCTTGATGATCCT 3901  
Db 4173 TGGGATATGATACAGACTAACTTTTGTGCTGCTTTTGTGCTTTTATCTTGATGATCCT 4230  
QY 3902 CAAGGAATACACCTAG 3919  
Db 4231 CAAGGAATACACCTAG 4248  
RESULT 9  
AX069342  
LOCUS AX069342  
DEFINITION Sequence 6 from Patent WO0102600.  
ACCESSION AX069342  
VERSION AX069342.1 GI:12579205  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Yuan, C.S.  
TITLE Detection of analytes using attenuated enzymes  
JOURNAL Patent: WO 0102600-A 6 11-JAN-2001;  
GENERAL ATOMICS (US)  
FEATURES  
Location/Qualifiers  
1. 7224  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
395..4192  
/notes="unnamed protein product; Human methionine synthase"  
CDS  
/codon\_start=1  
/protein\_id="CAC27300.1"  
/db\_xref="GI:12579206"  
/translation="MSPALQDLSQPEGLKTKTLRDEINAILQKRIWLDGGMQIMQRE  
KMBEHPRGQFKDHPKRNNDILSITPDVIYQIHKYLLAGADIIETNTFSYS  
IAQADYGLHAYRMMNCSAGVAKAEVTLQTKRFRVAGALPTNKTLSVPSVE  
RPDNYITFDELVEAOQAKGLDGVILLIETIPDANAKAALPALQNLFEKTA  
PRPTFISGTVKSGRTLSGTGGFVSVSHGELICGLNCALGAAMPFFIIGK  
CTTAYVLCYPNAGLPNTFGYDTPSNMAKHLDPAMDGLVNIWGGCCSTPDHRI  
AEAVNCKPRVPATAFEGHMLLSGLPEFPFRIPTNFVNIGERCNVAGSRKPAKLNA  
GNTEALCVAKVQVEMGAQVLDVNMDDGLDGPSTFRNLIASEPDIAPKVPCLIDS  
SNFAVBAGLKCCQKCVNSISLKEGDDFLKARKIKKYGAAMVWVAFDEGQATE  
TDTRIRVCTRAYHLLVKLGFPNDIIFDPNLTITGTMBEHNLYAINFTHATKVIKE  
TLPGARISGSLNLSPSFRGMEATREAHGVLVYHAIKSGMDMGI VNAAGLNVYDIIH  
KELLQLEDILMNKDPATEKLLRYAQTGCTGGKKVLOTDEWRNGPVEERLEYALVKG  
LKHIIIDTEARLNQKYPRLNIIEGPLMNGMKIVGDLFGAGKMFQVVKARVM  
KKAAGHILPFMEKEBEETRVNGVEEDPTQGTIVLATVAGDHDIGKNIVGVVLC  
NNFRVIDLGMVTPCDKILKAALDHADIIGLSGLITPSLDEMI FVAKEMERLAI RPL  
LIGADTSKTHATAKAPRYAPVHVLDSKSVVCSQLIDENLKDYEFEIMEEYE  
DIRQDHVESLKERRYLP LSOARKSGFQMDLSEPHVPKPTFGTQFVEDYDLOKLDVY  
IDWPKFPDQWLGKYPNRPKI FNDKTVGGEARKVYDDAHNMLNTLSQKLRARG  
VWGFPAQSLQDDIHLVABAAVPOAAPIATFFYGLRQQAEDKSDSTPEYCLSFAP  
LHSGYRISQDLGFLFVACFGVELSKAYEDDDSSIMVKALGDRLAFABELHERV  
RELWAYCGSQLDVAIDRLRYKIRPAPGYPQDPEKLTMMRLADI EQSGIRLT  
ESLAMAPASAVSLGYSFNLSKYPVAVGKISKDQVEDYALRNKNI SVAEVEKWLGPILGY  
DTD"

ORIGIN

Query Match 99.1%; Score 3883.2; DB 6; Length 7224;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3903; Conservative 0; Mismatches 13; Indels 2; Gaps 1;  
QY 2 GTCACTGTGAGAGCAGCTTCTTCTGCGCGCCCTCGCGAAGGAGAGCTCGACA 61  
Db 333 GTCACTGTGAGAGCAGCTTCTTCTGCGCGCCCTCGCGAAGGAGAGCTCGACA 392  
QY 62 ACATGTCAACCGCGCTCCAGACCTGTGCGAACCCGAGGCTCTGAAGAAAACCTCGCGG 121



Db 393 ACATGTACCCGCGCTCCAGACCTGTGCGAACCCGAGAGTCTGAAGAAAACCCCTGCGGG 452  
Qy 122 ATGAGATCAATGCGCAATCTGCGAAGAGGATATATGCTGCTGCGATGCGAGGATGCGGACCA 181  
Db 453 ATGAGATCAATGCGCAATCTGCGAAGAGGATATATGCTGCTGCGATGCGAGGATGCGGACCA 512  
Qy 182 TGATCCAGCGGAGAGCTAAACGAGAACACCTCCGAGGTCAGGAATTTAAAGATCATG 241  
Db 513 TGATCCAGCGGAGAGCTAAACGAGAACACCTCCGAGGTCAGGAATTTAAAGATCATG 572  
Qy 242 CCAGCGCTCAAGAGCAACATGACATTTTAAAGTATACCTCAGCTGATGTCATTTACC 301  
Db 573 CCAGCGCTCAAGAGCAACATGACATTTTAAAGTATACCTCAGCTGATGTCATTTACC 632  
Qy 302 AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAACAAATACTTTTAGCA 361  
Db 633 AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAACAAATACTTTTAGCA 692  
Qy 362 GCATAGTATTTGCCAGAGCTGACTATGGCCCTTGAACTCTGGCCCTACCGGATGAACATGT 421  
Db 693 GCATAGTATTTGCCAGAGCTGACTATGGCCCTTGAACTCTGGCCCTACCGGATGAACATGT 752  
Qy 422 GCTCTGAGGAGTGGCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTGAGGAGTGGCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGTGGCAGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG 541  
Db 813 GGTTTGTGGCAGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG 872  
Qy 542 AAAGCGGATATATAGGAACATCATTTGATGAGCTTTGTAAGACATACCAAGCAGG 601  
Db 873 AAAGCGGATATATAGGAACATCATTTGATGAGCTTTGTAAGACATACCAAGCAGG 932  
Qy 602 CCAAGGACTTCTGGATGGCGGCTTGATATCTTACTCATTTGAACATTTTGTGATCTG 661  
Db 933 CCAAGGACTTCTGGATGGCGGCTTGATATCTTACTCATTTGAACATTTTGTGATCTG 992  
Qy 662 CCAATGCCAGCGAGCTTGTGTGACCTCCAAATCTTTTGGAGAGAAATATGCTCCCC 721  
Db 993 CCAATGCCAGCGAGCTTGTGTGACCTCCAAATCTTTTGGAGAGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTCAGGAGCGATCGTGTGATAAAGTGGCGGACTCTTTCCGGACAGA 781  
Db 1053 GGCCTATCTTTATTCAGGAGCGATCGTGTGATAAAGTGGCGGACTCTTTCCGGACAGA 1112  
Qy 782 CAGGAGGGGATTTGTCTACGCTGTCTCATGAGAACCACTCTGCATTTGGATTAATTT 841  
Db 1113 CAGGAGGGGATTTGTCTACGCTGTCTCATGAGAACCACTCTGCATTTGGATTAATTT 1172  
Qy 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAATAATTTGGAATAATGTACACAG 901  
Db 1173 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAATAATTTGGAATAATGTACACAG 1232  
Qy 902 CCTATGCTCTGTATCCCAATGAGGTCCTCCCAACACCTTTGGTGAATATGATGAAA 961  
Db 1233 CCTATGCTCTGTATCCCAATGAGGTCCTCCCAACACCTTTGGTGAATATGATGAAA 1292  
Qy 962 CGCCTTCTATGATGGCCAGCACCTTAAGGATTTTGTCTATGGATGGCTTGGTCAATATAG 1021  
Db 1293 CGCCTTCTATGATGGCCAGCACCTTAAGGATTTTGTCTATGGATGGCTTGGTCAATATAG 1352  
Qy 1022 TTGGAGGATGCTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGGAGGATGCTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTTGAAGCTAGAGTTCACCTGCTGCTCTTTGAAGGACATATGTTACTGTCTGTC 1141  
Db 1413 ATTTGAAGCTAGAGTTCACCTGCTGCTCTTTGAAGGACATATGTTACTGTCTGTC 1472  
Qy 1142 TAGAGCCCTTCAGGATGGACCGTACCAACTTTGTGTTAACTTTGGAGAGCGCTGTAAATG 1201

Db 1473 TAGAGCCCTTCAGATTTGGACCGTACACCAACTTTTGTGTTAACTTTGGAGAGCGCTGTAAATG 1532  
Qy 1202 TTGCAGGATCAAGAAAGTTTGTCTAAACTCATCATGCGAGAAACTATGAAGAACCTTTGT 1261  
Db 1533 TTGCAGGATCAAGAAAGTTTGTCTAAACTCATCATGCGAGAAACTATGAAGAACCTTTGT 1592  
Qy 1262 GTGTTGCCAAAGTCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAAATGAGATGATG 1321  
Db 1593 GTGTTGCCAAAGTCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAAATGAGATGATG 1652  
Qy 1322 GCATGCTAGATGCTCAAGTGCATGACACAGATTTTGCAACTTTAAATGCTTCCGAGCCAG 1381  
Db 1653 GCATGCTAGATGCTCAAGTGCATGACACAGATTTTGCAACTTTAAATGCTTCCGAGCCAG 1712  
Qy 1382 ACATGCGAAAGGTACTTTGTGATGCACTCTCTCCAAATTTTGTGCTGATGAAGCTGGGT 1441  
Db 1713 ACATGCGAAAGGTACTTTGTGATGCACTCTCTCCAAATTTTGTGCTGATGAAGCTGGGT 1772  
Qy 1442 TAAAGTCTGCCAAGGAGTGCATTTGTCTCAATAGCATTAGTCTCAAGGAGGAGAGACG 1501  
Db 1773 TAAAGTCTGCCAAGGAGTGCATTTGTCTCAATAGCATTAGTCTCAAGGAGGAGAGACG 1852  
Qy 1502 ACTTCTTGGAGAAAGGCCAGGAAGATTAAAAAGTATGAGCTGCTTATGCTGCTCATGCTT 1561  
Db 1833 ACTTCTTGGAGAAAGGCCAGGAAGATTAAAAAGTATGAGCTGCTTATGCTGCTCATGCTT 1892  
Qy 1562 TTGATGAAGAAAGGACAGGCAACAGAACACAGACACAAAAATCAGAGTGTGACCCGGGCT 1621  
Db 1893 TTGATGAAGAAAGGACAGGCAACAGAACACAGACACAAAAATCAGAGTGTGACCCGGGCT 1952  
Qy 1622 ACCATGCTGCTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1681  
Db 1953 ACCATGCTGCTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 2012  
Qy 1682 TCCTTAACCATTTGGAGCTGGAAATGGAGAACACAACTTTGATGCCATTAATTTTATCCATG 1741  
Db 2013 TCCTTAACCATTTGGAGCTGGAAATGGAGAACACAACTTTGATGCCATTAATTTTATCCATG 2072  
Qy 1742 CAACAAAGTCACTTAAAGAAACATTTACTCGAGCCAGAAATAGTGGAGGCTTTTCCAACT 1801  
Db 2073 CAACAAAGTCACTTAAAGAAACATTTACTCGAGCCAGAAATAGTGGAGGCTTTTCCAACT 2132  
Qy 1802 TGTCTTCTCTCTCCGAGGAAATGGAAGCAATTCGAGAGCAATGCAATGGGGTTCCTTT 1861  
Db 2133 TGTCTTCTCTCTCCGAGGAAATGGAAGCAATTCGAGAGCAATGCAATGGGGTTCCTTT 2192  
Qy 1862 ACCATCAATCAAGTCTGGCATGACATGAGATAGTGAATGCTGGAAACCTCCCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGGCATGACATGAGGATAGTGAATGCTGGAAACCTCCCTGTGT 2252  
Qy 1922 ATGATGATATCCATTAAGAACTTCTTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db 2253 ATGATGATATCCATTAAGAACTTCTGCACTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
Qy 1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGAGGAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGAGGAGAAAG 2372  
Qy 2042 TCATTACAGCTGATGAGTGGAGAAATGGCCCTGCTGAGAGAACGCTTTGATGTCGCCCTTG 2101  
Db 2373 TCATTACAGCTGATGAGTGGAGAAATGGCCCTGCTGAGAGAACGCTTTGATGTCGCCCTTG 2432  
Qy 2102 TGAAGGGCATTTGAAAAACATTTATTAGGATATCTGAGGAGCCAGGTAAACCAAAAAA 2161  
Db 2433 TGAAGGGCATTTGAAAAACATTTATTAGGATATCTGAGGAGCCAGGTAAACCAAAAAA 2492  
Qy 2162 AATATCCCGACCTCTCAATATATTAATGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
Db 2493 AATATCCCGACCTCTCAATATATTAATGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATTAAGTCAAGCCCGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATTAAGTCAAGCCCGGTTA 2612

```
QY 2282 TGAAGAAGCGTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAAACCCAGAG 2341
Db 2613 TGAAGAAGCGTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAAACCCAGAG 2672
QY 2342 TGCCTTAAACGGCAGCTAGAGAAGAGGACCCCTTACCAGGCGACCATCTGCTGGCCACTG 2401
Db 2673 TGCCTTAAACGGCAGCTAGAGAAGAGGACCCCTTACCAGGCGACCATCTGCTGGCCACTG 2732
QY 2402 TTAAGAAGCGAGCTGCAGACATAGGCAAGAAACATAGTTGGAGTAGTCTCTGGCTGCAATA 2461
Db 2733 TTAAGAAGCGAGCTGCAGACATAGGCAAGAAACATAGTTGGAGTAGTCTCTGGCTGCAATA 2792
QY 2462 ATTTCCGAGTTATTTGATTTAGAGTTCATGACTCCATGTGATAAGATACCTGAAGCTGCTC 2521
Db 2793 ATTTCCGAGTTATTTGATTTAGAGTTCATGACTCCATGTGATAAGATACCTGAAGCTGCTC 2852
QY 2522 TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTTCCCTGGATGAAA 2581
Db 2853 TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTTCCCTGGATGAAA 2912
QY 2582 TGATTTTGTGTCGAAGGAATGGAGAGATAGCTATTAAGGATTCATCTGTTGATGGAG 2641
Db 2913 TGATTTTGTGTCGAAGGAATGGAGAGATAGCTATTAAGGATTCATCTGTTGATGGAG 2972
QY 2642 GAGCAACCACTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701
Db 2973 GAGCAACCACTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 3032
QY 2702 TAACTCATGTCTCGACGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 2761
Db 3033 TAACTCATGTCTCGACGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 3092
QY 2762 ATCTAAAGATGAATACTTTGAGGAATCATGGAAAGATATGAAGATATTAGACAGGACC 2821
Db 3093 ATCTAAAGATGAATACTTTGAGGAATCATGGAAAGATATGAAGATATTAGACAGGACC 3152
QY 2822 ATTATGACTCTCAAGGAGGAGATACTTACCCTTAAGTCAAGCCAGAAAGTGGTT 2881
Db 3153 ATTATGACTCTCAAGGAGGAGATACTTACCCTTAAGTCAAGCCAGAAAGTGGTT 3212
QY 2882 TCCAAATGGATTGGTGTCTGAACCTCACCCAGTGAAGCCACGCTTTATTTGGACCCAGG 2941
Db 3213 TCCAAATGGATTGGTGTCTGAACCTCACCCAGTGAAGCCACGCTTTATTTGGACCCAGG 3272
QY 2942 TCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTGACTTGAAGCCTTCTTTG 3001
Db 3273 TCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTGACTTGAAGCCTTCTTTG 3332
QY 3002 ATGCTGGCAGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTAAAGACA 3061
Db 3333 ATGCTGGCAGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTAAAGACA 3392
QY 3062 AAACAGTAGTGTGAGAGGCCAGGAAGTCTACGATGATGCCACCAATATGCTGAACACAC 3121
Db 3393 AAACAGTAGTGTGAGAGGCCAGGAAGTCTACGATGATGCCACCAATATGCTGAACACAC 3452
QY 3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGTTCCTGGCCAGCACAGATA 3181
Db 3453 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGTTCCTGGCCAGCACAGATA 3512
QY 3182 TCCAAAGCAGATTCACCTGTACGACAGGCTGTGCTGTGCCAGGCTGCAGAGCCCATAG 3241
Db 3513 TCCAAAGCAGATTCACCTGTACGAGGCTGTGCTGTGCCAGGCTGCAGAGCCCATAG 3572
QY 3242 CCACCTTCTATGGTTAAGCCAAAGCTGGAAGGACTCTCCAGCAGCAGGACCATACT 3301
Db 3573 CCACCTTCTATGGTTAAGCCAAAGCTGGAAGGACTCTCCAGCAGCAGGACCATACT 3632
QY 3302 ACTGCTCTCAGACTTTCAGCTTCCCTTGCATTTCTGGCATCCGTGACTACTTGGGCGCTGT 3361
Db 3633 ACTGCTCTCAGACTTTCAGCTTCCCTTGCATTTCTGGCATCCGTGACTACTTGGGCGCTGT 3692
```

```
QY 3362 TTGCGGTGCTGCTTTGGGTAGAAAGAGCTGAGCAAGGCGCTATGAGGATGATGCTGACG 3421
Db 3693 TTGCGGTGCTGCTTTGGGTAGAAAGAGCTGAGCAAGGCGCTATGAGGATGATGCTGACG 3752
QY 3422 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCCGGCTGGCGAGAGCCCTTTGCAAGAG 3481
Db 3753 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCCGGCTGGCGAGAGCCCTTTGCAAGAG 3812
QY 3482 AGCTCCATGAAGAAGTTCCGCGAGAACTGTGGGCGCTACTTGGCAGTGAGCAGCTGACG 3541
Db 3813 AGCTCCATGAAGAAGTTCCGCGAGAACTGTGGGCGCTACTTGGCAGTGAGCAGCTGACG 3872
QY 3542 TCGCAGACCTGCGAAGGTTGCGGTACAAAGGCGCATCCGCCCGGCTCTCTGGCTACCCAGCC 3601
Db 3873 TCGCAGACCTGCGCAGGCTGGGTACAAAGGCGCATCCGCCCGGCTCTCTGGCTACCCAGCC 3932
QY 3602 AGCCCGACCAACCGAAGCTCACATGTGGAGACTTCGCGAGACATCGAGCAGTCTACAG 3661
Db 3933 AGCCCGACCAACCGAAGCTCACATGTGGAGACTTCGCGAGACATCGAGCAGTCTACAG 3992
QY 3662 GCATTAGGTTAACAGAAATCATTAGCAATGGCAGCTGCTTCCAGCAGTCTCAGGCGCTCTACT 3721
Db 3993 GCATTAGGTTAACAGAAATCATTAGCAATGGCAGCTGCTTCCAGCAGTCTCAGGCGCTCTACT 4052
QY 3722 TCTCCAATTTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAAGGATCAGGTTGAGG 3781
Db 4053 TCTCCAATTTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAAGGATCAGGTTGAGG 4112
QY 3782 ATTATGCAATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCTTTGGACCCATTT 3841
Db 4113 ATTATGCAATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCTTTGGACCCATTT 4172
QY 3842 TGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTTCCTTTTATCTTGATGATCCT 3901
Db 4173 TGGATATGATACAGACTAACT--TTTTTTTTTTTTTTTTCCTTTTATCTTGATGATCCT 4230
QY 3902 CAAGGAAATACAACCTAG 3919
Db 4231 CAAGGAAATACAACCTAG 4248
```

## RESULT 10

HSU73338 7224 bp mRNA linear PRI 02-JAN-1997  
LOCUS Human methionine synthase mRNA, complete cds.

DEFINITION U73338  
ACCESSION U73338  
VERSION U73338.1 GI:1763268

KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 7224)  
AUTHORS Chen, L.H., Liu, M.-L., Hwang, H.-Y., Chen, L.-S., Korenberg, J. and

Shane, B.

TITLE Human methionine synthase: cDNA cloning, gene localization and

expression

JOURNAL J. Biol. Chem. (1996) In press

REFERENCE 2 (bases 1 to 7224)

AUTHORS Chen, L.H., Liu, M.-L., Hwang, H.-Y., Chen, L.-S., Korenberg, J. and

Shane, B.

TITLE Direct Submission

JOURNAL Submitted (03-OCT-1996) Nutritional Sciences, University of

California, 117 Morgan Hall, Berkeley, CA 94720-3104, USA

FEATURES  
Location/Qualifiers

1..7224

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="1"

/map="1q42.3-1q43"

/tissue\_type="liver"

CDS	/dev_stage="fetus"	362	GCACTAGTATTGCCAGCTGACTATGGCCCTTGAACACTTGGCTACCGGATGAACATGT	Qy
	/395_..4192			
	/EC_number="2.1.1.13"	693	GCACTAGTATTGCCAGCTGACTATGGCCCTTGAACACTTGGCTACCGGATGAACATGT	Db
	/codon_start=1			
	/product="methionine synthase"	422	GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGAGAGTAACTCTCCAGACAGGAATTAAGA	Qy
	/protein_id="AAB39704.1"			
	/db_xref="GI:1763269"	753	GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGAGAGTAACTCTCCAGACAGGAATTAAGA	Db
	/translation="MSPALQDLSQPEGLKKTLRDEINAILQKRIWLDGNGMTIQRE	482	GGTTTGTGGCAGGGGCTCTGGGTCGGAATAAAGACACTCTCTGTGTGCCCATCTGTGG	Qy
	KLNEHFRGQFDFKDRAPLKNNDILSIQPDVYIQHKEYLLAGADIIETNTPSSYS			
	LAADYGLEHLAYRMNCSAGVARKAAEEVLTQIRKRFVAGALPNTKLSVSPVS	813	GGTTTGTGGCAGGGGCTCTGGGTCGGAATAAAGACACTCTCTGTGTGCCCATCTGTGG	Db
	PRPIFTSGLTIDELVEAYQOAGLLDGVILLIETIPTDANAKAALFALQNLFEKTA			
	PPFIISGTIDVSKSGRLSQGTGEGFVSVSHGPELIGLNCALGAEMEPFTEIIGK	542	AAAGGCCGGAATATAGAAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG	Qy
	CTTAVLCYNAGLPNTFGDYDETPSMMAKHLKDFAMDGLVNIWGVGCGSTPDHIREI			
	AEAVNCKPVPATPAEGHMLLSGLEPFRIGPYTNFNTGERCNVAGSRKPAKLINA	873	AAAGGCCGGAATATAGAAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG	Db
	SNFAVIEAGLCCGCKCIVNSISLKEGEDDFLEKARKIKKYGAAMVMAFDEBQATE			
	TDTRIKTRYALHLVKGLGFNPNDIIFDPNIIITGTGMBEHLIYAINFTHATKVIKE	602	CCTATGCTCTGTTGGATGGCGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG	Qy
	TLPGARISGLLSNLFSPFRGMEAIIEAMHGVFLYHAIKSGMDMGIVNAGNLPVYDIIH			
	KELLQLEDLILNNDPEATEKLRYACTQGTGKKVLOTDEWRNGPVEERLEVALVKG	933	CCTATGCTCTGTTGGATGGCGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG	Db
	LEKHLIIEDETEARLNQKYPRLNIIISGPLMNGMKIVGDLFGAGKMLPQVIKSARVM			
	KAVGHILIPMEKEEREETRVNGTVEBEDPYQGTIVLATVKGVDVHDIGKNIWGVVLC	662	CCTATGCTCTGTTGGATGGCGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG	Qy
	NNFVIDLGVMTCDKILKAALDHKADIIGLSLITSLDEMI FVAKEMERLAIRIPL			
	LIGATTSTKHTATVIAAPRYSAPIVHVLDAKS VVVCQSLLDENLKDVEEYETIWERPE	721	CCAAATGCCAAGGAGCGCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC	Db
	DIRODHVESLKERLYPLNSQARKSFQMDMLSEBPKVPTF IGTOVEDYDLOKLVDY			
	IDWKPFDVQLRKYRNRPFKLPNDKTVGGEARVYDDAHNMLNLIISQKLRLARG	1053	GGCTTATCTTTTATTTCCAGGACGATCGTGTGATAAAAGTGGCGGACTCTTTTCCGACAGA	Qy
	VGVFWPAQS QDDIHLTAERAAVPOAEPPIATF IGLRQOAKSDASTEPYICLSDFIAP			
	LHSGIRLYGLFAVACFGVEELS KAYEDDGDYSSIMVKALGRDLAFAPAEELHVR	782	CAGGAGAGGATTTGTTCATCAGCGTGTCTCATCGAGAACCACTCTGCTATTTGGATTAATTT	Db
	RELWAYCGSEQLDRLRLRYKIRPAPGYPSQPDHTEKLTMWRLADIQSTGIRLT			
	ESLAMAPASAVSGLYFSLNLSKYFAVGKISKQVEDYALRNKINSVAEVEKWLGPILGY	1113	CAGGAGAGGATTTGTTCATCAGCGTGTCTCATCGAGAACCACTCTGCTATTTGGATTAATTT	Qy
	DTD"			
variation	/note="polymorphism G1004T; creates an amino acid change of Ala to Ser"			
	/replace="t"			
variation	/note="polymorphism G1158A; creates an amino acid change of Cys to Tyr at amino acid 255"			
	/replace="a"			
variation	/note="polymorphism A3150G; creates an amino acid change of Asp to Gly at amino acid 919"			
	/replace="g"			
	7184..7189			
polyA_signal				
ORIGIN				
	Query Match 99.1%; Score 3883.2; DB 8; Length 7224;			
	Best Local Similarity 99.6%; Pred. No. 0;			
	Matches 3903; Conservative 0; Mismatches 13; Indels 2; Gaps 1;			
Qy	2 GTACCTGTGGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	1081	TTGGAGGATGCTGTGGGTCAACACCAGATCATATCAGGGAAATTTGCTGGAAGCTGTGAAA	Qy
Db	333 GTACCTGTGGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA			
Qy	62 ACATGTCAACCGCGCTCCAGACCTGTGCAACCCGAGGTCTGAAGAAAACCTCGCGG	1353	TTGGAGGATGCTGTGGGTCAACACCAGATCATATCAGGGAAATTTGCTGGAAGCTGTGAAA	Db
Db	393 ACATGTCAACCGCGCTCCAGACCTGTGCAACCCGAGGTCTGAAGAAAACCTCGCGG			
Qy	122 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTATGTTGCTGATGGAGGATGGGACCA	1082	ATTGTAAGCTTAGAGTTCCACCTGCGCATCTGCTTTTGAAGGACATATGTTACTGTGGTC	Qy
Db	453 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTATGTTGCTGATGGAGGATGGGACCA			
Qy	182 TGATCCAGCGGAGAGCTAAACGAAGAACCTTTCCGAGGTTCAGGAATTTTAAAGATCATG	1413	ATTGTAAGCTTAGAGTTCCACCTGCGCATCTGCTTTTGAAGGACATATGTTACTGTGGTC	Db
Db	513 TGATCCAGCGGAGAGCTAAACGAAGAACCTTTCCGAGGTTCAGGAATTTTAAAGATCATG			
Qy	242 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1201	TAGAGCGCTTCAGGATTTGGACCGTACACCAACTTTGTTAACTTTGGAGCGCTGTAAATG	Qy
Db	301 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	573 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1473	TAGAGCGCTTCAGGATTTGGACCGTACACCAACTTTGTTAACTTTGGAGCGCTGTAAATG	Db
Db	361 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1202	TTGCAGGATCAAGGAAAGTTTGTCTAACTCATCTGTCAGGAGAACTATGAAGAGACCTTGT	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1532	TTGCAGGATCAAGGAAAGTTTGTCTAACTCATCTGTCAGGAGAACTATGAAGAGACCTTGT	Db
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1261	TTGCAGGATCAAGGAAAGTTTGTCTAACTCATCTGTCAGGAGAACTATGAAGAGACCTTGT	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1592	TTGCAGGATCAAGGAAAGTTTGTCTAACTCATCTGTCAGGAGAACTATGAAGAGACCTTGT	Db
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1321	GTGTTGCCAAAGTGCAGGTGGAAATGGAGCCCGAGGTGTTGGATGTCAACATGAGATCATG	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1593	GTGTTGCCAAAGTGCAGGTGGAAATGGAGCCCGAGGTGTTGGATGTCAACATGAGATCATG	Db
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1381	GCATGCTAGATGCTCCAGTGCATGACACGAGTTTGGCACTTAATTTGCTTCCGAGCCAG	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1712	GCATGCTAGATGCTCCAGTGCATGACACGAGTTTGGCACTTAATTTGCTTCCGAGCCAG	Db
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1441	ACATCGCAAGAGTACCTTTTGTGATCGACTCTCCCAATTTTGTGATTTGATTTGAGAGCTGGT	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1772	ACATCGCAAGAGTACCTTTTGTGATCGACTCTCCCAATTTTGTGATTTGATTTGAGAGCTGGT	Db
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			
Qy	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG	1501	TAAAGTGTGCTGCCAAGGAGAGTGCATTTGTCAATAGCATTTAGTCTGAGGAGAGGAGACG	Qy
Db	692 CCAGCGCGTGAAGGACCAATGACATTTTAAAGTATACTCAGCTTCAGGAATTTTAAAGATCATG			

1773	Db	TAAAGTGTCTCCAAAGGAGTGCATTTGTCAATAGCAATTAGTCTGAAGGAAGAGAGGACG	1832
1502	Qy	ACTTCTTGGAGAGGCGCAGGAAGATTAAAGATPATGGAGCTGCTATGGTGGTCAATGGCTTT	1561
1833	Db	ACTTCTTGGAGAGGCGCAGGAAGATTAAAGATPATGGAGCTGCTATGGTGGTCAATGGCTTT	1892
1562	Qy	TTGATGAAGAAGACAGGCGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCT	1621
1893	Db	TTGATGAAGAAGACAGGCGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCT	1952
1622	Qy	ACCATCTGCTTGTGAAAAAATCTGGGCTTTAAATCCAAATGACATATATTTTGTGACCCTAATA	1681
1953	Db	ACCATCTGCTTGTGAAAAAATCTGGGCTTTAAATCCAAATGACATATATTTTGTGACCCTAATA	2012
1682	Qy	TCTTAACCATTTGGGCTGGGATGGAGGACACAACTTGTATGCCATTAATTTTATCCCATG	1741
2013	Db	TCTTAACCATTTGGGCTGGGATGGAGGACACAACTTGTATGCCATTAATTTTATCCCATG	2072
1742	Qy	CAACAAAGTCAATTAAGAAACATTAACCTGGAGCAGAAATAAGTGGAGTCTTCCAACT	1801
2073	Db	CAACAAAGTCAATTAAGAAACATTAACCTGGAGCAGAAATAAGTGGAGTCTTCCAACT	2132
1802	Qy	TGTCCTTCTCCTTCCGAGGAATGGAAGCCATTGCGAAGCAATGCAATGGGGTTTCCCTTT	1861
2133	Db	TGTCCTTCTCCTTCCGAGGAATGGAAGCCATTGCGAAGCAATGCAATGGGGTTTCCCTTT	2192
1862	Qy	ACCATGCATCAAGTCTGGCATGGACATGGAGATGTAAGTGGAAACCTCCCTGTGT	1921
2193	Db	ACCATGCATCAAGTCTGGCATGGACATGGAGATGTAAGTGGAAACCTCCCTGTGT	2252
1922	Qy	ATGATGATATCCCATAGGAATCTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCC	1981
2253	Db	ATGATGATATCCCATAGGAATCTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCC	2312
1982	Qy	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGSCACAGAGGGAAGAAAG	2041
2313	Db	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGSCACAGAGGGAAGAAAG	2372
2042	Qy	TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCTTGAGTATGCCCTTG	2101
2373	Db	TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCTTGAGTATGCCCTTG	2432
2102	Qy	TGAAGGSCATTGAAACATATTTATGAGATACTGAGGAAGCAGGTTAAACCAAAAAA	2161
2433	Db	TGAAGGSCATTGAAACATATTTATGAGATACTGAGGAAGCAGGTTAAACCAAAAAA	2492
2162	Qy	AATATCCCGACCTCTCAATATATTAATTAAGGACCCCTGATGAATGGAATGAAAAATGTTTG	2221
2493	Db	AATATCCCGACCTCTCAATATATTAATTAAGGACCCCTGATGAATGGAATGAAAAATGTTTG	2552
2222	Qy	GTGATCTTTTGGAGCTGGAAAAATTTGTTCTACTCAGGTTATAAGTCAAGCCCGGTTA	2281
2553	Db	GTGATCTTTTGGAGCTGGAAAAATTTGTTCTACTCAGGTTATAAGTCAAGCCCGGTTA	2612
2282	Qy	TGAAGAGGCTGTGGCCACTTATCCCTTTCATGGAAAGAAAGAGAGAAACACAGAG	2341
2613	Db	TGAAGAGGCTGTGGCCACTTATCCCTTTCATGGAAAGAAAGAGAGAAACACAGAG	2672
2342	Qy	TGCTTTAAACGGCAGATGAGAAAGAGGACCCCTTACCAGGGCACCATCGTGCCTGGCACCTG	2401
2673	Db	TGCTTTAAACGGCAGATGAGAAAGAGGACCCCTTACCAGGGCACCATCGTGCCTGGCACCTG	2732
2402	Qy	TTAAAGGCGACGTGCACGACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA	2461
2733	Db	TTAAAGGCGACGTGCACGACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA	2792
2462	Qy	ATTTCCGAGTTATTTGATTTAGGAGTCAATGATCCATGCTGATAGATCACTGGAAGCTGCTC	2521
2793	Db	ATTTCCGAGTTATTTGATTTAGGAGTCAATGATCCATGCTGATAGATCACTGGAAGCTGCTC	2852
2522	Qy	TTGACCAAAAGCAGATATTAATTTGGCCGTGCAGGACTCATCTCCTTCCCTGGATGAA	2581

Qy 3662 GCATTAGGTTAAACAGATCATTAGCAATGGCACTGCTTACAGAGTCTCAGGCGCTCTACT 3721  
Db 3993 GCATTAGGTTAAACAGATCATTAGCAATGGCACTGCTTACAGAGTCTCAGGCGCTCTACT 4052  
Qy 3722 TCTCCAAATTTGCAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGCAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 4112  
Qy 3782 ATTATGCAATGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
Db 4113 ATTATGCAATGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4172  
Qy 3842 TGGGATATGATACAGACTAATTTTGTGCTGAGGTTGAGAAATGGCTTGGATGATGATCCT 3901  
Db 4173 TGGGATATGATACAGACTAAT - TTTTGTGCTGAGGTTGAGAAATGGCTTGGATGATGATCCT 4230  
Qy 3902 CAAGGAATAACAACCTAG 3919  
Db 4231 CAAGGAATAACAACCTAG 4248

RESULT 11  
AR300095 LOCUS AR300095 7224 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 1 from patent US 6537759.  
ACCESSION AR300095  
VERSION AR300095.1 GI:31687382  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 7224)  
AUTHORS Stanton, V. P. Jr.  
TITLE Polyglutamate synthetase gene sequence variances having utility in determining the treatment of disease  
JOURNAL Patent: US 6537759-A 1 25-MAR-2003;  
VARIATION Variagenics, Inc.; Cambridge, MA  
FEATURES  
source 1. 7224  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 98.9%; Score 3876.2; DB 6; Length 7224;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3896; Conservative 0; Mismatches 20; Indels 2; Gaps 1;  
Qy 2 GTACCTGTGAGAGCAGCTCTTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61  
Db 333 GTACCTGTGAGAGCAGCTCTTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392  
Qy 62 ACATGTCAACCGCGCTCCAGACCTGTCGCAACCGGAGGTCGCAAGAAACCTGCGGG 121  
Db 393 ACATGTCAACCGCGCTCCAGACCTGTCGCAACCGGAGGTCGCAAGAAACCTGCGGG 452  
Qy 122 ATGAGATCAATGCCATCTTCAGAAAGAGGATATGTTGCTGCTGATGAGGAGGATGGGACCA 181  
Db 453 ATGAGATCAATGCCATCTTCAGAAAGAGGATATGTTGCTGCTGATGAGGAGGATGGGACCA 512  
Qy 182 TGATCCAGCGGAGAGCTTAAACGAAGAACACTTCCGAGGTGAGGAAATTTAAAGATCATG 241  
Db 513 TGATCCAGCGGAGAGCTTAAACGAAGAACACTTCCGAGGTGAGGAAATTTAAAGATCATG 572  
Qy 242 CCAGCGCGCTGAAAGGCAACATGACATTTTAAAGTATTAACCTCAGCGCTGATGTCATTACC 301  
Db 573 CCAGCGCGCTGAAAGGCAACATGACATTTTAAAGTATTAACCTCAGCGCTGATGTCATTACC 632  
Qy 302 AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 361  
Db 633 AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 692  
Qy 362 GCACTAGTATTGCCAAGCTGACTATGGCCCTTGAACACTTGGCCCTACCGGATGAACATGT 421

Db 693 GCACTAGTATTGCCAAGCTGACTATGGCCCTTGAACACTTGGCCCTACCGGATGAACATGT 752  
Qy 422 GCTCTGCAGAGTGGCCAGAAAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTGCAGAGTGGCCAGAAAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAATTAAGACACTCTCTGTGTCCCATCTCTGG 541  
Db 813 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAATTAAGACACTCTCTGTGTCCCATCTCTGG 872  
Qy 542 AAAGGCCGATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
Db 873 AAAGGCCGATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 932  
Qy 602 CCAAGGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG 661  
Db 933 CCAAGGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG 992  
Qy 662 CCAATGCCAAGGAGCCTTGTGCTGACTCCTCAAAATCTTTTGGAGGAGAAATATGCTCCC 721  
Db 993 CCAATGCCAAGGAGCCTTGTGCTGACTCCTCAAAATCTTTTGGAGGAGAAATATGCTCCC 1052  
Qy 722 GGCCTATCTTTTTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCCGACAGA 781  
Db 1053 GGCCTATCTTTTTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCCGACAGA 1112  
Qy 782 CAGGAGAGGATTTGTTCATCAGCGTGTCTCATGAGGAAACCACTCTGCTGATTTGATTAAT 841  
Db 1113 CAGGAGAGGATTTGTTCATCAGCGTGTCTCATGAGGAAACCACTCTGCTGATTTGATTAAT 1172  
Qy 842 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATTTGTAACAAG 901  
Db 1173 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATTTGTAACAAG 1232  
Qy 902 CCTATGCTCTGTTATCCCAATGACAGTCTTCCCAACACTTTGGTGAATGATGATAA 961  
Db 1233 CCTATGCTCTGTTATCCCAATGACAGTCTTCCCAACACTTTGGTGAATGATGATAA 1292  
Qy 962 CGCCTTCTATGATGGCCCAAGCACCTAAAGGATTTTGTCTGATGGTGGCTTGGTCAATATAG 1021  
Db 1293 CGCCTTCTATGATGGCCCAAGCACCTAAAGGATTTTGTCTGATGGTGGCTTGGTCAATATAG 1352  
Qy 1022 TTGGAGGATCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGGAGGATCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGCTC 1141  
Db 1413 ATTGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGCTC 1472  
Qy 1142 TAGAGCCCTTCAGGATTTGGACCGGTACACCAACTTTTGTAACTTTGGAGAGCGCTGTAATG 1201  
Db 1473 TAGAGCCCTTCAGGATTTGGACCGGTACACCAACTTTTGTAACTTTGGAGAGCGCTGTAATG 1532  
Qy 1202 TTGCAGGATCAAGGAAGTTTGTAACTCATATGCGAGGAACCTATGAAGAACCTTGT 1261  
Db 1533 TTGCAGGATCAAGGAAGTTTGTAACTCATATGCGAGGAACCTATGAAGAACCTTGT 1592  
Qy 1262 GTGTTGCAAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGATGATG 1321  
Db 1593 GTGTTGCAAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGATGATG 1652  
Qy 1322 GCATGCTAGATGTTCCAAAGTGCATGACAGATTTTGCACACTTAAATTTCCGAGCCAG 1381  
Db 1653 GCATGCTAGATGTTCCAAAGTGCATGACAGATTTTGCACACTTAAATTTCCGAGCCAG 1712  
Qy 1382 ACATCGAAAGGTACCTTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGGT 1441  
Db 1713 ACATCGAAAGGTACCTTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGGT 1772  
Qy 1442 TAAAGTGTCTGCCAAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1501



Db	1773	TAAAGTGTGTCACAAAGGAAGTGCAATTGTCAATAGCATTTAGTCTCAAGAAAGGAGGAGACG	1832
Qy	1502	ACTTCTTGGAGAGGCGCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTGGTCAATGGCTT	1561
Db	1833	ACTTCTTGGAGAGGCGCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTGGTCAATGGCTT	1892
Qy	1562	TTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCAACCCGGGCGCT	1621
Db	1893	TTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCAACCCGGGCGCT	1952
Qy	1622	ACCATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATTTATTTTGACCCCTTAATA	1681
Db	1953	ACCATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATTTATTTTGACCCCTTAATA	2012
Qy	1682	TCCTAAACATTTGGGACTGGATGGAGGAACACAACTTGTAAGCCATTAAATTTTATCCATG	1741
Db	2013	TCCTAAACATTTGGGACTGGATGGAGGAACACAACTTGTAAGCCATTAAATTTTATCCATG	2072
Qy	1742	CAACAAAAAGTCATTAAGAAAAATTTACCTGGAGCCAGAAATTAAGTGGAGGTCTTTCCAACT	1801
Db	2073	CAACAAAAAGTCATTAAGAAAAATTTACCTGGAGCCAGAAATTAAGTGGAGGTCTTTCCAACT	2132
Qy	1802	TGTCCTTCTCCTTCGAGGAATGGAAGCCATTTCGAGAAAGCAATGATGGGGTTTTCCTTT	1861
Db	2133	TGTCCTTCTCCTTCGAGGAATGGAAGCCATTTCGAGAAAGCAATGATGGGGTTTTCCTTT	2192
Qy	1862	ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT	1921
Db	2193	ACCATGCAATCAAGTCTGGCATGGACATGGGGATAGTGAATGCTGGAAACCTCCCTGTGT	2252
Qy	1922	ATGATGATATCCATAAGGAATCTTCGACAGTCTGTGAAGATCTCATCTGGAAATAAGACC	1981
Db	2253	ATGATGATATCCATAAGGAATCTTCGACAGTCTGTGAAGATCTCATCTGGAAATAAGACC	2312
Qy	1982	CTGAGGCCACTGAGAAGCTCTTAACGTTATGCCAGACTCAAGGCCACAGAGGGAGAAAG	2041
Db	2313	CTGAGGCCACTGAGAAGCTCTTAACGTTATGCCAGACTCAAGGCCACAGAGGGAGAAAG	2372
Qy	2042	TCATTCAGACTGATGAGTGGGAATGCGCTGTGGAAGACGCTTCAGATATGCCCTTG	2101
Db	2373	TCATTCAGACTGATGAGTGGGAATGCGCTGTGGAAGACGCTTCAGATATGCCCTTG	2432
Qy	2102	TGAAGGGCATTGAAAAACATATTTAGAGGATCTTGAGGAAGCCAGGTTAAACCAAAAAA	2161
Db	2433	TGAAGGGCATTGAAAAACATATTTAGAGGATCTTGAGGAAGCCAGGTTAAACCAAAAAA	2492
Qy	2162	AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTG	2221
Db	2493	AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTG	2552
Qy	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTA	2281
Db	2553	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTA	2612
Qy	2282	TGAAGAGGCTGTGCGCACCTTATCCCTTTCATGCAAAAAAGAAAGAAACACAGAG	2341
Db	2613	TGAAGAGGCTGTGCGCACCTTATCCCTTTTCATGCAAAAAAGAAAGAAACACAGAG	2672
Qy	2342	TGCTTTAAACGGCACAGTAGAAGAGAGGACCCCTTACCAGGGCACCACTCGTGTGGCCACTG	2401
Db	2673	TGCTTTAAACGGCACAGTAGAAGAGAGGACCCCTTACCAGGGCACCACTCGTGTGGCCACTG	2732
Qy	2402	TTAAAGGGCAGCTGCACGACATAGGCAAGAACATATGTTGGAGTAGTCTTGGCTGCAATA	2461
Db	2733	TTAAAGGGCAGCTGCACGACATAGGCAAGAACATATGTTGGAGTAGTCTTGGCTGCAATA	2792
Qy	2462	ATTTCCGAGTTATGATTTAGAGATCATGACTCCATGTGATTAAGATACCTGAAAGCTGCTC	2521
Db	2793	ATTTCCGAGTTATGATTTAGAGATCATGACTCCATGTGATTAAGATACCTGAAAGCTGCTC	2852
Qy	2522	TTGACCAACAAGCAGATATAATTTGGCTGTGAGGACTCATCATCTCTTCCCTCGGATGAAA	2581
Db	2853	TTGACCAACAAGCAGATATAATTTGGCTGTGAGGACTCATCATCTCTTCCCTCGGATGAAA	2912
Qy	2582	TGATTTTGTGTCACAAAGGAATCGAGAGATTAGCTATAAGGATTCCAATTGTTGATTGGAG	2641
Db	2913	TGATTTTGTGTCACAAAGGAATCGAGAGATTAGCTATAAGGATTCCAATTGTTGATTGGAG	2972
Qy	2642	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTG	2701
Db	2973	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTG	3032
Qy	2702	TAAATCCATGTCTCTGGAGCGCTCCAAGAGTGTGTGGTGTGTCTCCAGCTGTAGATGAAA	2761
Db	3033	TAAATCCATGTCTCTGGAGCGCTCCAAGAGTGTGTGGTGTGTCTCCAGCTGTAGATGAAA	3092
Qy	2762	ATCTAAAGAGTGAATTAATTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC	2821
Db	3093	ATCTAAAGAGTGAATTAATTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC	3152
Qy	2822	ATTATGATGCTCTCAAGGAGAGGAGATACCTTACCTTAACTCAAGCCAGAAAAAGTGGTT	2881
Db	3153	ATTATGATGCTCTCAAGGAGAGGAGATACCTTACCTTAACTCAAGCCAGAAAAANTNGTT	3212
Qy	2882	TCCAAAATGGAATGCTGTCTGAACTCCACCCAGTGAAGCCCAAGTTTATTTGGGACCCAGG	2941
Db	3213	TCCAAAATGGAATGCTGTCTGAACTCCACCCAGTGAAGCCCAAGTTTATTTGGGACCCAGG	3272
Qy	2942	TCCTTTGAAGACTATGACTTCGAGAAGCTGGTGGACTACATTGACTTGAAGCCCTTCTTTTG	3001
Db	3273	TCCTTTGAAGACTATGACTTCGAGAAGCTGGTGGACTACATTGACTTGAAGCCCTTCTTTTG	3332
Qy	3002	ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTCCCGCAAGATATTTAAACGACA	3061
Db	3333	ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTTCCCAAGATATTTAAACGACA	3392
Qy	3062	AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCCAAAATGCTGGAACACAC	3121
Db	3393	AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCCAAAATGCTGGAACACAC	3452
Qy	3122	TGATTTAGTCAAAAAGAAAACTCCGGGCCCGGGGTGTGGTGTGGTCTTGCCACAGCAGTA	3181
Db	3453	TGATTTAGTCAAAAAGAAAACTCCGGGCCCGGGGTGTGGTGTGGTCTTGCCACAGCAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACCGCGAGGCTGTGTGTGCCCGCAGGCTGCGAGGCCATAG	3241
Db	3513	TCCAAGACGACATTCACCTGTACCGCGAGGCTGTGTGTGCCCGCAGGCTGCGAGGCCATAG	3572
Qy	3242	CCACTTTCTATGGGTTAAAGCAACAGGCTGAGAGGACTCTGCCAGCACGAGCCATACT	3301
Db	3573	CCACTTTCTATGGGTTAAAGCAACAGGCTGAGAGGACTCTGCCAGCACGAGCCATACT	3632
Qy	3302	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTTGACTTACCTGGGCCCTGT	3361
Db	3633	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTTGACTTACCTGGGCCCTGT	3692
Qy	3362	TTGCCGTTGCTGCTTTTGGGTTAGAAGAGCTGACAAAGGCTCTATGAGATGATGGTGACG	3421
Db	3693	TTGCCGTTGCTGCTTTTGGGTTAGAAGAGCTGACAAAGGCTCTATGAGATGATGGTGACG	3752
Qy	3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCCGCTGGCAGAGGCCCTTTGCAAGAG	3481
Db	3753	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCCGCTGGCAGAGGCCCTTTGCAAGAG	3812
Qy	3482	AGCTCCATGAAGAGTTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG	3541
Db	3813	AGCTCCATGAAGAGTTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG	3872
Qy	3542	TCGACAGACTTCGGAAGGTTGCGGTACAAGGCACTCCGCGGGCTCTGGCTACCCAGCC	3601
Db	3873	TCGACAGACTTCGCGAGGCTGCGGTACAAGGCACTCCGCGGGCTCTGGCTACCCAGCC	3932
Qy	3602	AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTCGAGACATCGAGCAGTCTACAG	3661
Db	3933	AGCCCGACCAACCGAGAGGCTCACCATGTGGAGACTTCACATGTGGAGACTTCGACAGATCGAGCAGTCTACAG	3992

[illegible]

Qy	1502	ACTTCTTGGAGAGCCAGGAGAGTTAAAAAGTATGGAGCTGCTATGCTGCTCATGCGTT	1561
Dd	1833	ACTTCTTGGAGAGCCAGGAGAGTTAAAAAGTATGGAGCTGCTATGCTGCTCATGCGTT	1892
Qy	1562	TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAGAAATCAGAGTGTGCAACCGGCGCT	1621
Dd	1893	TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAGAAATCAGAGTGTGCAACCGGCGCT	1952
Qy	1622	ACCATCTGCTTGTGAAAGAACTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA	1681
Dd	1953	ACCATCTGCTTGTGAAAGAACTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA	2012
Qy	1682	TCCTAAACCATTTGGGACTGGAAATGAGGAACACAACTTGTATGCCATTAATTTTATCCATG	1741
Dd	2013	TCCTAAACCATTTGGGACTGGAAATGAGGAACACAACTTGTATGCCATTAATTTTATCCATG	2072
Qy	1742	CAACAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGTCTTTCCAACT	1801
Dd	2073	CAACAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGTCTTTCCAACT	2132
Qy	1802	TGTCCTTCTCCTCCGAGGAATGGAAGCCATTCGAGAAAGCAATGCAATGGGCTTTTCCCTT	1861
Dd	2133	TGTCCTTCTCCTCCGAGGAATGGAAGCCATTCGAGAAAGCAATGCAATGGGCTTTTCCCTT	2192
Qy	1862	ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCTCTGT	1921
Dd	2193	ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCTCTGT	2252
Qy	1922	ATGATGATATCCATAAGGAACTTTGCGAGCTCTGTGAAGATCTCATCTGGAATTAAGACC	1981
Dd	2253	ATGATGATATCCATAAGGAACTTTGCGAGCTCTGTGAAGATCTCATCTGGAATTAAGACC	2312
Qy	1982	CTGAGGCCACTGAGAGCTCTTAAGTTATGCCAGACTCAAGGCAACAGAGGGAAGAAAG	2041
Dd	2313	CTGAGGCCACTGAGAGCTCTTAAGTTATGCCAGACTCAAGGCAACAGAGGGAAGAAAG	2372
Qy	2042	TCATTCAGACTGATGAGTGGAGAAATGCGCCCTGTGGAAGAGCGCTTGATGATGCCCTTG	2101
Dd	2373	TCATTCAGACTGATGAGTGGAGAAATGCGCCCTGTGGAAGAGCGCTTGATGATGCCCTTG	2432
Qy	2102	TGAAGGCAATTTGAAAGCAATATTTAGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA	2161
Dd	2433	TGAAGGCAATTTGAAAGCAATATTTAGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA	2492
Qy	2162	AATATCCCGACCTCTCAATATAATGAAAGAACCCCTGATGAATGGAATGAATTTGTTG	2221
Dd	2493	AATATCCCGACCTCTCAATATAATGAAAGAACCCCTGATGAATGGAATGAATTTGTTG	2552
Qy	2222	GTGATCTTTTGGAGCTGGAAAGATTTTCTACCTCAGGTTATAAGTTCAGCCCGGTTA	2281
Dd	2553	GTGATCTTTTGGAGCTGGAAAGATTTTCTACCTCAGGTTATAAGTTCAGCCCGGTTA	2612
Qy	2282	TGAAGAGGCTGTGGCCACCTTATCCCTTTTCATGAAAAAGAAAGAAAGAACCCAGAG	2341
Dd	2613	TGAAGAGGCTGTGGCCACCTTATCCCTTTTCATGAAAAAGAAAGAAAGAACCCAGAG	2672
Qy	2342	TGCTTAAACGACAGTGAAGAGGAGCCCTTACAGGGCAACATCGTGTGCGCACTG	2401
Dd	2673	TGCTTAAACGACAGTGAAGAGGAGCCCTTACAGGGCAACATCGTGTGCGCACTG	2732
Qy	2402	TTAAAGGCGAGTGCACACATAGGCAAGACATAGTTGGAGTAGTCCCTTGGCTGCAATA	2461
Dd	2733	TTAAAGGCGAGTGCACACATAGGCAAGACATAGTTGGAGTAGTCCCTTGGCTGCAATA	2792
Qy	2462	ATTTCCGAGTTAATGATTTAGAGTCACTGATCCATGTGATAAGATCTGAAAGCTGCTC	2521
Dd	2793	ATTTCCGAGTTAATGATTTAGAGTCACTGATCCATGTGATAAGATCTGAAAGCTGCTC	2852
Qy	2522	TTGACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCACTCCTTCCCTGAGTAA	2581
Dd	2853	TTGACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCACTCCTTCCCTGAGTAA	2912

Qy	2582	TGATTTTGTTCGCAAGAAATGGAGAGATTAGCTATAAGGATTTCCATTGTGATTGGAG	2641
Dd	2913	TGATTTTGTTCGCAAGAAATGGAGAGATTAGCTATAAGGATTTCCATTGTGATTGGAG	2972
Qy	2642	GAGCAACCACTTCAAAAACCCACACAGAGCTTAAAAATAGCTTCGAGATACAGTGCACTG	2701
Dd	2973	GAGCAACCACTTCAAAAACCCACACAGAGCTTAAAAATAGCTTCGAGATACAGTGCACTG	3032
Qy	2702	TAAATCCATGTCCTGGAGCGCTCCAGAGTGTGGTGTGCTTCCAGCTGTAGATGAAA	2761
Dd	3033	TAAATCCATGTCCTGGAGCGCTCCAGAGTGTGGTGTGCTTCCAGCTGTAGATGAAA	3092
Qy	2762	ATCTAAAGGATGAATACCTTTGAGGAATCATGGAAGATATGAAGATATTAAGACAGACC	2821
Dd	3093	ATCTAAAGGATGAATACCTTTGAGGAATCATGGAAGATATGAAGATATTAAGACAGACC	3152
Qy	2822	ATTATGAGTCTCTCAAGGAGAGGATATTAACCTTTAAGTCAAGCCAGAAAAAGTGGTT	2881
Dd	3153	ATTATGAGTCTCTCAAGGAGAGGATATTAACCTTTAAGTCAAGCCAGAAAAANTNGTT	3212
Qy	2882	TCCAAATGGAATGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG	2941
Dd	3213	TCCAAATGGAATGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG	3272
Qy	2942	TCCTTGAAGACTATGACTCGCAGAGCTGGTGGACTACATTGACTGGAAGCCTTTCTTTTG	3001
Dd	3273	TCCTTGAAGACTATGACTCGCAGAGCTGGTGGACTACATTGACTGGAAGCCTTTCTTTTG	3332
Qy	3002	ATGTCCTGGAGCTCGGGGCAAGTACCCGAATCGAGGCTTTCCCAAGATATTTAAACACA	3061
Dd	3333	ATGTCCTGGAGCTCGGGGCAAGTACCCGAATCGAGGCTTTCCCAAGATATTTAAACACA	3392
Qy	3062	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGATGCCCAACATATGCTGACACAC	3121
Dd	3393	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGATGCCCAACATATGCTGACACAC	3452
Qy	3122	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGGTGGTGGTGTGCTGCGCCAGACAGAGTA	3181
Dd	3453	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGGTGGTGGTGTGCTGCGCCAGACAGAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACGACAGAGCTGCTGTGCCCCAGGCTGACAGAGCCATAG	3241
Dd	3513	TCCAAGACGACATTCACCTGTACGACAGAGCTGCTGTGCCCCAGGCTGACAGAGCCATAG	3572
Qy	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAAGGACTCTGCCAGCACCGGAGCCATCT	3301
Dd	3573	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAAGGACTCTGCCAGCACCGGAGCCATCT	3632
Qy	3302	ACTGCCCTCAGACCTTCATCGCTCCCTTGCATTCCTGGCATCCGTCATCTACCTGGCCCTGT	3361
Dd	3633	ACTGCCCTCAGACCTTCATCGCTCCCTTGCATTCCTGGCATCCGTCATCTACCTGGCCCTGT	3692
Qy	3362	TTGCCGTTGCTGCTTTGGGTTAGAAAGCTGAGCAAGGCTTATGAGGATGATGGTGAAG	3421
Dd	3693	TTGCCGTTGCTGCTTTGGGTTAGAAAGCTGAGCAAGGCTTATGAGGATGATGGTGAAG	3752
Qy	3422	ACTACAGAGCATCATGGTCAAGGCGCTGGGGGACCGCTGGGAGAGGCTTTTGCAGAG	3481
Dd	3753	ACTACAGAGCATCATGGTCAAGGCGCTGGGGGACCGCTGGGAGAGGCTTTTGCAGAG	3812
Qy	3482	AGCTCCATGAAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGAAG	3541
Dd	3813	AGCTCCATGAAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGAAG	3872
Qy	3542	TCGACAGCCTGCGAAGGTTGCGGTACAAAGGCACTCCGCCCGGCTCCTGGCTACCCAGCC	3601
Dd	3873	TCGACAGCCTGCGAAGGTTGCGGTACAAAGGCACTCCGCCCGGCTCCTGGCTACCCAGCC	3932
Qy	3602	AGCCCGACACACCGAGAGGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3661
Dd	3933	AGCCCGACACACCGAGAGGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3992
Qy	3662	GCATTAGGTTAAACAGAAATCATTAGCAATGGCACCTGCTTTCAGCAGTCTCAGGCTCTACT	3721



```
Db 3993 GCATTAGGTTACAGAAATCATTAGCAATGGCACTGCTTTCAGAGCTCTCAGGCTCTACT 4052
Qy 3722 TCTCAAATTTGAAGTCCAAATPATTTTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 3781
Db 4053 TCTCAAATTTGAAGTCCAAATPATTTTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 4112
Qy 3782 ATTATGATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCGTTGGACCCATTTT 3841
Db 4113 ATTATGATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCGTTGGACCCATTT 4172
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTATCTTTGATGATCCT 3901
Db 4173 TGGGATATGATACAGACTAAC - TTTTTTTTTTTTTTTTGGCTTTTTTATCTTTGATGATCCT 4230
Qy 3902 CAAGGAATCAACCTAG 3919
Db 4231 CAAGGAATCAACCTAG 4248

RESULT 13
AR562003
LOCUS AR562003 7224 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759200.
ACCESSION AR562003
VERSION AR562003.1 GI:53975655
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7224)
AUTHORS Stanton, V. P. Jr.
TITLE Thymidine phosphorylase gene sequence variances having utility in
determining the treatment of disease
JOURNAL Patent: US 6759200-A 1 06-JUL-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source 1. 7224
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 98.9%; Score 3876.2; DB 6; Length 7224;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3896; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

Qy 2 GTCACTGTGGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61
Db 333 GTCACTGTGGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392
Qy 62 ACATGTCAACCGCGCTCCAAGACCTGTGCAACCCGAGGCTCTGAAGAAACCTCGCGG 121
Db 393 ACATGTCAACCGCGCTCCAAGACCTGTGCAACCCGAGGCTCTGAAGAAACCTCGCGG 452
Qy 122 ATGAGATCAATGCAATTCCTGCAAGAGGATTAATGCTGCTGATGAGGAGGATGGGACCA 181
Db 453 ATGAGATCAATGCAATTCCTGCAAGAGGATTAATGCTGCTGATGAGGAGGATGGGACCA 512
Qy 182 TGATCCAGCGGAGAGCTAAACCAAGCAACACTTCCGAGGCTCAGGAATTTAAAGATCATG 241
Db 513 TGATCCAGCGGAGAGCTAAACCAAGCAACACTTCCGAGGCTCAGGAATTTAAAGATCATG 572
Qy 242 CCAGCGCGCTGAAGGCAACAAATGACATTTTAACTATAACTCAGCCTGATGTCATTTACC 301
Db 573 CCAGCGCGCTGAAGGCAACAAATGACATTTTAACTATAACTCAGCCTGATGTCATTTACC 632
Qy 302 AAATCCATAAGGAATATCTGTGCTGGGCGAGATATCAATTTGAAACAAATATTTTAGCA 361
Db 633 AAATCCATAAGGAATATCTGTGCTGGGCGAGATATCAATTTGAAACAAATATTTTAGCA 692
Qy 362 GCATAGTATTTGCCAAGCTGACTATGCGCTTGAACACTTGGCCCTACCGGATGACATGT 421
Db 693 GCATAGTATTTGCCAAGCTGACTATGCGCTTGAACACTTGGCCCTACCGGATGACATGT 752
```

```
Qy 422 GCTCTGCGAGGAGTGGCCAGAAAAGCTGCGCAGAGGAGTAACTCTCCAGACAGAGAAATTAAGA 481
Db 753 GCTCTGCGAGGAGTGGCCAGAAAAGCTGCGCAGAGGAGTAACTCTCCAGACAGAGAAATTAAGA 812
Qy 482 GGTTTGTGGCAGGCGGCTCTGGGTCCGACTAATAAGACACTCTCTGTCTCCCATCTCTGG 541
Db 813 GGTTTGTGGCAGGCGGCTCTGGGTCCGACTAATAAGACACTCTCTGTCTCCCATCTCTGG 872
Qy 542 AAAGGCGCGAATTATAGGAAACATCACTTTGATGAGCTTGTGTAAGCATATCAAGAGCAGG 601
Db 873 AAAGGCGCGAATTATAGGAAACATCACTTTGATGAGCTTGTGTAAGCATATCAAGAGCAGG 932
Qy 602 CAAAAGACTTCTGGATGGCGGGGTGATATCTTACTCATTTGAAACTATTTTGTGATCTG 661
Db 933 CAAAAGACTTCTGGATGGCGGGGTGATATCTTACTCATTTGAAACTATTTTGTGATCTG 992
Qy 662 CCAATGCCAAGGCGAGCTTCTTTGCACTCCAAATCTTTTGTAGGAGAAATATGCTCCCC 721
Db 993 CCAATGCCAAGGCGAGCTTCTTTGCACTCCAAATCTTTTGTAGGAGAAATATGCTCCCC 1052
Qy 722 GGCTATCTTTTATTTCAAGGACGATCGTTGATAAAAGTGGGCGGACTCTTTTCCGACAGA 781
Db 1053 GGCTATCTTTTATTTCAAGGACGATCGTTGATAAAAGTGGGCGGACTCTTTTCCGACAGA 1112
Qy 782 CAGGAGAGGATTTGTCATCAGCGTGTCTCATGAGAGAACCACTCTGCAATTGGATTAATTT 841
Db 1113 CAGGAGAGGATTTGTCATCAGCGTGTCTCATGAGAGAACCACTCTGCAATTGGATTAATTT 1172
Qy 842 GTGCTTTGGGTGCGAGCTGAGATGAGACCTTTTATGAAATAATTTGGAATAATGTACACAG 901
Db 1173 GTGCTTTGGGTGCGAGCTGAGATGAGACCTTTTATGAAATAATTTGGAATAATGTACACAG 1232
Qy 902 CCTATGCTCTCTGTATCCCAATGCGAGCTCTCCCAACACCTTTTGTGACTATGATGAAA 961
Db 1233 CCTATGCTCTCTGTATCCNAATGCGAGTCTTCCCAACACCTTTTGTGACTATGATGAAA 1292
Qy 962 CGCTTCTATGATGGCCAAAGCACTAAAGGATTTTGTCTATGAGATGGCTTGGTCAATATAG 1021
Db 1293 CGCTTCTATGATGGCCAAAGCACTAAAGGATTTTGTCTATGAGATGGCTTGGTCAATATAG 1352
Qy 1022 TTGAGGATGCTCTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081
Db 1353 TTGAGGATGCTCTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412
Qy 1082 ATTGTAAGCTTAGAGTTCCACCTGCGACATCTTTTGAAGGACATATGCTTACTGCTGGTC 1141
Db 1413 ATTGTAAGCTTAGAGTTCCACCTGCGACATCTTTTGAAGGACATATGCTTACTGCTGGTC 1472
Qy 1142 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGTAACTATGAGAGCGCTGTAATG 1201
Db 1473 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGTAACTATGAGAGCGCTGTAATG 1532
Qy 1202 TTGAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAAACTATGAAGAGCCTTGT 1261
Db 1533 TTGAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAAACTATGAAGAGCCTTGT 1592
Qy 1262 GTCTTGCAGGAGTGGGAAATGGGAGCCAGGCTTGGATGTCACCAATGATGATG 1321
Db 1593 GTCTTGCAGGAGTGGGAAATGGGAGCCAGGCTTGGATGTCACCAATGATGATG 1652
Qy 1322 GCATGCTAGATGTTCCAAAGTCAATGACCAAGTATTTTGCATTTTGTGCTGATTTGAGCTGGT 1381
Db 1653 GCATGCTAGATGTTCCAAAGTCAATGACCAAGTATTTTGCATTTTGTGCTGATTTGAGCTGGT 1712
Qy 1382 ACATCGAAGAGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAGCTGGT 1441
Db 1713 ACATCGAAGAGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAGCTGGT 1772
Qy 1442 TAAAGTGTCCCAAGGAGTGCATTTGTCATTAAGTATGCTTGAAGGAGGAGGACG 1501
Db 1773 TAAAGTGTCCCAAGGAGTGCATTTGTCATTAAGTATGCTTGAAGGAGGAGGACG 1832
```

Qy	1502	ACTCTTGGAGAAAGCCGAGAAATTA	1561	TTGATGAAGAGGACAGGCAACAGAA
Db	1833	ACTTTCTGGAGAAAGCCGAGAAATTA	1892	TTGATGAAGAGGACAGGCAACAGAA
Qy	1562	TTGATGAAGAGGACAGGCAACAGAA	1621	TTGATGAAGAGGACAGGCAACAGAA
Db	1893	TTGATGAAGAGGACAGGCAACAGAA	1952	TTGATGAAGAGGACAGGCAACAGAA
Qy	1622	ACCATCTGCTTGTGAAAAA	1681	ACCATCTGCTTGTGAAAAA
Db	1953	ACCATCTGCTTGTGAAAAA	2012	ACCATCTGCTTGTGAAAAA
Qy	1682	TCCTAAACCATTTGGGAACTGGGCTT	1741	TCCTAAACCATTTGGGAACTGGGCTT
Db	2013	TCCTAAACCATTTGGGAACTGGGCTT	2072	TCCTAAACCATTTGGGAACTGGGCTT
Qy	1742	CAACAAAAGTCATTAAAGAA	1801	CAACAAAAGTCATTAAAGAA
Db	2073	CAACAAAAGTCATTAAAGAA	2132	CAACAAAAGTCATTAAAGAA
Qy	1802	TGTCCTTCTCTCTCCGAGGAATGGA	1861	TGTCCTTCTCTCTCCGAGGAATGGA
Db	2133	TGTCCTTCTCTCTCCGAGGAATGGA	2192	TGTCCTTCTCTCTCCGAGGAATGGA
Qy	1862	ACCATGCAATCAAGTCTGGCATG	1921	ACCATGCAATCAAGTCTGGCATG
Db	2193	ACCATGCAATCAAGTCTGGCATG	2252	ACCATGCAATCAAGTCTGGCATG
Qy	1922	ATGATGATATCCATAAGGAATCT	1981	ATGATGATATCCATAAGGAATCT
Db	2253	ATGATGATATCCATAAGGAATCT	2312	ATGATGATATCCATAAGGAATCT
Qy	1982	CTGAGGCCACTTGAGAAGCTCTTA	2041	CTGAGGCCACTTGAGAAGCTCTTA
Db	2313	CTGAGGCCACTTGAGAAGCTCTTA	2372	CTGAGGCCACTTGAGAAGCTCTTA
Qy	2042	TCATTTCAGACTGATGAGTGGAGA	2101	TCATTTCAGACTGATGAGTGGAGA
Db	2373	TCATTTCAGACTGATGAGTGGAGA	2432	TCATTTCAGACTGATGAGTGGAGA
Qy	2102	TGAAGGGCATTTGAAAAACATATTA	2161	TGAAGGGCATTTGAAAAACATATTA
Db	2433	TGAAGGGCATTTGAAAAACATATTA	2492	TGAAGGGCATTTGAAAAACATATTA
Qy	2162	AATATCCCGACCTCTCAATATTA	2221	AATATCCCGACCTCTCAATATTA
Db	2493	AATATCCCGACCTCTCAATATTA	2552	AATATCCCGACCTCTCAATATTA
Qy	2222	GTGATCTTTTTGGAGCTGGAAAA	2281	GTGATCTTTTTGGAGCTGGAAAA
Db	2553	GTGATCTTTTTGGAGCTGGAAAA	2612	GTGATCTTTTTGGAGCTGGAAAA
Qy	2282	TGAAGAAAGCTGTGTGGCCACCTTA	2341	TGAAGAAAGCTGTGTGGCCACCTTA
Db	2613	TGAAGAAAGCTGTGTGGCCACCTTA	2672	TGAAGAAAGCTGTGTGGCCACCTTA
Qy	2342	TGCTTAAACGGCACAGTATAGAA	2401	TGCTTAAACGGCACAGTATAGAA
Db	2673	TGCTTAAACGGCACAGTATAGAA	2732	TGCTTAAACGGCACAGTATAGAA
Qy	2402	TTAAAGGGCAGCTGCACGACATAG	2461	TTAAAGGGCAGCTGCACGACATAG
Db	2733	TTAAAGGGCAGCTGCACGACATAG	2792	TTAAAGGGCAGCTGCACGACATAG
Qy	2462	ATTTCCGAGTTTATTAATTTAGG	2521	ATTTCCGAGTTTATTAATTTAGG
Db	2793	ATTTCCGAGTTTATTAATTTAGG	2852	ATTTCCGAGTTTATTAATTTAGG
Qy	2522	TTGACCAACAAAGCAGATATTA	2581	TTGACCAACAAAGCAGATATTA
Db	2853	TTGACCAACAAAGCAGATATTA	2912	TTGACCAACAAAGCAGATATTA
Qy	2582	TGATTTTTGTGTGCGCAAGAAAT	2641	TGATTTTTGTGTGCGCAAGAAAT

Db	2913	TGATTTTGTGTCAAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTGTTGATTGGAG	2972
Qy	2642	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2973	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACCTG	3032
Qy	2702	TAAATCATGTCTCGGACGGGTCCAGAGTGTGGTGTGTTCCCAAGCTGTTAGATGAAA	2761
Db	3033	TAATCATGTCTCGGACGGGTCCAGAGTGTGGTGTGTTCCCAAGCTGTTAGATGAAA	3092
Qy	2762	ATCTAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGACC	2821
Db	3093	ATCTAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGACC	3152
Qy	2822	ATTATGAGTCTCTCAAGGAGAGAGATACTTACCTTTAAGTCAAGCCAGAAAAAGTGTT	2881
Db	3153	ATTATGAGTCTCTCAAGGAGAGAGATACTTACCTTTAAGTCAAGCCAGAAAAANTGTT	3212
Qy	2882	TCCAAATGGATTTGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG	2941
Db	3213	TCCAAATGGATTTGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG	3272
Qy	2942	TCTTTGAGACTATGACCTGCAGAAAGCTGGTGAAGCTACATTTGACTGGAAAGCTTTCTTTG	3001
Db	3273	TCTTTGAGACTATGACCTGCAGAAAGCTGGTGAAGCTACATTTGACTGGAAAGCTTTCTTTG	3332
Qy	3002	ATGCTCTGGCAGCTCCCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA	3061
Db	3333	ATGCTCTGGCAGCTCCCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA	3392
Qy	3062	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCCAAAATATGCTGAACACAC	3121
Db	3393	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCCAAAATATGCTGAACACAC	3452
Qy	3122	TGATTAGTCAAAAGAAACTCCGGGCCCGGGGTGTGTTGGGTTCTGGCCAGCACAGAGTA	3181
Db	3453	TGATTAGTCAAAAGAAACTCCGGGCCCGGGGTGTGTTGGGTTCTGGCCAGCACAGAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACGAGAGGCTGCTGTGCCCCAGGCTGCAGAGCCCATAG	3241
Db	3513	TCCAAGACGACATTCACCTGTACGAGAGGCTGCTGTGCCCCAGGCTGCAGAGCCCATAG	3572
Qy	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCACGGAGCATACT	3301
Db	3573	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCACGGAGCATACT	3632
Qy	3302	ACTGCCTCTCAGACTTCATCGCTCCCTTTCGATTCCTGGCATTCGTGACTACTGGGCCCTGT	3361
Db	3633	ACTGCCTCTCAGACTTCATCGCTCCCTTTCGATTCCTGGCATTCGTGACTACTGGGCCCTGT	3692
Qy	3362	TTCCGCTTGCTGCTTTGGGGTAGAGAGCTGAGCAAGGCTATGAGATGATGCTGCACG	3421
Db	3693	TTCCGCTTGCTGCTTTGGGGTAGAGAGCTGAGCAAGGCTATGAGATGATGCTGCACG	3752
Qy	3422	ACTACAGCAGCATCATGGTCAAGCGCTCGGGGACCGGCTGGCAGAGGCTTTTCAGAGAAG	3481
Db	3753	ACTACAGCAGCATCATGGTCAAGCGCTCGGGGACCGGCTGGCAGAGGCTTTTCAGAGAAG	3812
Qy	3482	AGCTCCAAGAAAGATTCCGCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG	3541
Db	3813	AGCTCCAAGAAAGATTCCGCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG	3872
Qy	3542	TCGCAGACCTCGAAGGTTGGGTACAGGGCATCCGCCCGGCTCTGGGCTACCCAGCC	3601
Db	3873	TCGCAGACCTCGCAGGCTCGGTACAGGGCATCCGCCCGGCTCTGGGCTACCCAGCC	3932
Qy	3602	AGCCCGACCAACCGAGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG	3661
Db	3933	AGCCCGACCAACCGAGAAGCTCACCATGTGGAGACTTCGAGACATCGAGCAGCTCTACAG	3992
Qy	3662	GCATTAGGTTAAAGAAATCATTTAGCAATGGCACCTGCTTACGAGTCTCAGGCCCTCTACT	3721

Db	3993	GCATTAGGTTAA	CAGAAATCAT	TAGCAATGGC	CACTGCTTCAGCAGTCTCAGGCGCTCTTACT	4053			
Qy	3722	TCTCCAAATTTGAAGTCCAAATATATTTGCTGTGGGGAAGATTTTCCAAAGGATCAGGTTGAGG	3781						
Db	4053	TCTCCAAATTTGAAGTCCAAATATATTTGCTGTGGGGAAGATTTTCCAAAGATCAGGTTGAGG	4112						
Qy	3782	ATTATGCATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCGTTGGACCCCATTT	3841						
Db	4113	ATTATGCATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCGTTGGACCCCATTT	4172						
Qy	3842	TGGGATATGATACAGACTTAACCTTTTTTTTTTTTTTTTTTTTGGCTTTTATCTTGATGATCCT	3901						
Db	4173	TGGGATATGATACAGACTTAAC--TTTTTTTTTTTTTTTGGCTTTTTTATCTTGATGATCCT	4230						
Qy	3902	CAAGGAAATACAACCTAG	3919						
Db	4231	CAAGGAAATACAACCTAG	4248						
RESULT 14									
LOCUS	AR482320		3856 bp	DNA	linear	PAT 14-MAY-2004			
DEFINITION	Sequence 75 from patent US 6703197.								
ACCESSION	AR482320								
VERSION	AR482320.1	GI:47244609							
KEYWORDS									
SOURCE		Unknown.							
ORGANISM		Unknown.							
REFERENCE	1 (bases 1 to 3856)								
AUTHORS	Gravel,R.A., Rozen,R., LeClerc,D., Goyette,P. and Campeau,B.								
TITLE	Human methionine synthase: cloning, and methods for evaluating risk								
JOURNAL	of neural tube defects, cardiovascular disease, and cancer								
	Patent: US 6703197-A 75 09-MAR-2004;								
	Martinez R&D, Inc.; Montreal;								
	CAN;								
FEATURES		Location/Qualifiers							
source	1. 3856								
	/organism="unknown"								
	/mol_type="genomic DNA"								
ORIGIN									
	Query Match	98.3%;	Score 3851;	DB 6;	Length 3856;				
	Best Local Similarity	99.9%;	Pred. No. 0;						
	Matches 3851;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
Qy	64	ATGTCAACCCGGCTCCAAAGACCTGTCCCAACCCGGAAGTCTGAAGAAAACCCCTGCGGGAT	123						
Db	1	ATGTCACCCCGCTCCAAAGACCTGTCCCAACCCGGAAGTCTGAAGAAAACCCCTGCGGGAT	60						
Qy	124	GAGATCAATGCATTTCTGCAGAGAGAGATTTATGCTGCTGGATGAGGAGATGGGGACCATG	183						
Db	61	GAGATCAATGCGATTTCTGCAGAGAGAGATTTATGCTGCTGGATGAGGAGATGGGGACCATG	120						
Qy	184	ATCCAGCGGGAGAGCTAAACGAGAACACTTCCGAGGTGAGGAATTTAAAGATCATGCC	243						
Db	121	ATCCAGCGGGAGAGCTAAACGAGAACACTTCCGAGGTGAGGAATTTAAAGATCATGCC	180						
Qy	244	AGGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCTATTACCAA	303						
Db	181	AGGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCTATTACCAA	240						
Qy	304	ATCCATAAGGAATACTTGTGCTGGCTGGGGCAGATATCAATGAAACAAATACCTTTTAGCAGC	363						
Db	241	ATCCATAAGGAATACTTGTGCTGGCTGGGGCAGATATCAATGAAACAAATACCTTTTAGCAGC	300						
Qy	364	ACTAGTATTGCCCAAGCTGACTATGCGCTTTGAAACACTTTGGCCTACCGGATGAACATGTGC	423						
Db	301	ACTAGTATTGCCCAAGCTGACTATGCGCTTTGAAACACTTTGGCCTACCGGATGAACATGTGC	360						
Qy	424	TCTGCAGAGTGGCCAGAAAAGCTGCGAGAGGTAACTCTCCAGACAGGAAATTAAGAGG	483						
Db	361	TCTGCAGAGTGGCCAGAAAAGCTGCGAGAGGTAACTCTCCAGACAGGAAATTAAGAGG	420						

Qy	1564	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGGCTTAC	1622
Db	1501	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGGCTTAC	1560
Qy	1624	CATCTGCTTGTGA AAAACCTGGGCTTTAAATCCAAATGACATTAATTTTGCACCCATAATC	1683
Db	1561	CATCTGCTTGTGA AAAACCTGGGCTTTAAATCCAAATGACATTAATTTTGCACCCATAATC	1620
Qy	1684	CTAACCATTTGGGACTGGAAATGGAGGAACACAACCTTGTATGCCATTAATTTATCCCATGCA	1743
Db	1621	CTAACCATTTGGGACTGGAAATGGAGGAACACAACCTTGTATGCCATTAATTTATCCCATGCA	1680
Qy	1744	ACAAAAGTCAATTAAGAAAACATTAACCTGGAGCCAGAATAAGTCGAGGTCTTTTCCAACTTG	1803
Db	1681	ACAAAAGTCAATTAAGAAAACATTAACCTGGAGCCAGAATAAGTCGAGGTCTTTTCCAACTTG	1740
Qy	1804	TCCTTCTCTTCGAGGAATGGAGCCATTCGAGAGGAACAATGCAATGGGGTTTTCCTTTTAC	1863
Db	1741	TCCTTCTCTTCGAGGAATGGAGCCATTCGAGAGGAACAATGCAATGGGGTTTTCCTTTTAC	1800
Qy	1864	CATGCAATCAAGTCTGCATCGGACATGGAGATAGTGAATCTCGGAAAACCTCCCTGTGTAT	1923
Db	1801	CATGCAATCAAGTCTGCATCGGACATGGAGATAGTGAATCTCGGAAAACCTCCCTGTGTAT	1860
Qy	1924	GATGATATCCATAAGGAACTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCT	1983
Db	1861	GATGATATCCATAAGGAACTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCT	1920
Qy	1984	GAGGCCACTGAGAAAGCTCTTAAGTTATGCCAGACTCMAGGCHACAGAGGGAAGAAGTC	2043
Db	1921	GAGGCCACTGAGAAAGCTCTTAAGTTATGCCAGACTCMAGGCHACAGAGGGAAGAAGTC	1980
Qy	2044	ATTCCAGACTGATGAGTGGAGAAATGGCCCTGTTCGAGAACCCCTTGAGTATGCCCTTGTG	2103
Db	1981	ATTCCAGACTGATGAGTGGAGAAATGGCCCTGTTCGAGAACCCCTTGAGTATGCCCTTGTG	2040
Qy	2104	AAGGGCATTTGAAAACAATATTTATGGAGTACTGAGGAAGCCAGGTTAAACCAAAAAAAA	2163
Db	2041	AAGGGCATTTGAAAACAATATTTATGGAGTACTGAGGAAGCCAGGTTAAACCAAAAAAAA	2100
Qy	2164	TATCCCGACCTCTCAATATATAATTGAAGGACCCCTGATGTAATGGAATGAAAATTTGTTGT	2223
Db	2101	TATCCCGACCTCTCAATATATAATTGAAGGACCCCTGATGTAATGGAATGAAAATTTGTTGT	2160
Qy	2224	GATCTTTTTCGAGCTCGAAAATGTTTTCTACCTCAGGTATAAGTCAGCCCCGGTTATG	2283
Db	2161	GATCTTTTTCGAGCTCGAAAATGTTTTCTACCTCAGGTATAAGTCAGCCCCGGTTATG	2220
Qy	2284	AAGAAGGCTGTGGCCACCTTTATCCCTTTTCATGGA AAAAGAGAAGAACCCAGAGTG	2343
Db	2221	AAGAAGGCTGTGGCCACCTTTATCCCTTTTCATGGA AAAAGAGAAGAACCCAGAGTG	2280
Qy	2344	CTTAACGGGCAAGTAGAAGAAAGAGGACCTTTACAGGGCAACCATCGTGTGGCCACTGTT	2403
Db	2281	CTTAACGGGCAAGTAGAAGAAAGAGGACCTTTACAGGGCAACCATCGTGTGGCCACTGTT	2340
Qy	2404	AAAGGCGAGCTGCAGCATATAGGCAAGACATAGTTCGAGTAGTCCCTTGCTGCAATAAT	2463
Db	2341	AAAGGCGAGCTGCAGCATATAGGCAAGACATAGTTCGAGTAGTCCCTTGCTGCAATAAT	2400
Qy	2464	TTCCGAGTTATTGATTTTAGGAGTCAATGACTCCATGTGTATTAAGATPACTGAAAGCTGCTCTT	2523
Db	2401	TTCCGAGTTATTGATTTTAGGAGTCAATGACTCCATGTGTATTAAGATPACTGAAAGCTGCTCTT	2460
Qy	2524	GACCACAAAGCAGATATAATTTGGCTGTFCAGGACTCATCTCTTCCCTTGGATGAAATG	2583
Db	2461	GACCACAAAGCAGATATAATTTGGCTGTFCAGGACTCATCTCTTCCCTTGGATGAAATG	2520
Qy	2584	ATTTTTTCTTCCAGGAAATGGAGAGATTAGCTATAAGGATTCCTATTTGATTGGAGGA	2643
Db	2521	ATTTTTTCTTCCAGGAAATGGAGAGATTAGCTATAAGGATTCCTATTTGATTGGAGGA	2580
Qy	2644	GCAACCACTTCAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTCGACCTGTAT	2703

Db	2581	GCACCACTTCAAAACCCACACAGCAGTTAAATAGCTCGAGATACAGTGCACTGTN	2640
Qy	2704	ATCCATGTCCTGGAGCGCTCAAGAGTGTGGTGGTGTTCCTCCAGCTGTTAGATGAAAAT	2763
Db	2641	NNCCAATGTCCTGGAGCGCTCAAGAGTGTGGTGGTGTTCCTCCAGCTGTTAGATGAAAAT	2700
Qy	2764	CTAAGGATGAATACTTTTGAAGAAATCATATGGAAGAAATATGAAGATATTAAGACAGGACCAT	2823
Db	2701	CTAAGGATGAATACTTTTGAAGAAATCATATGGAAGAAATATGAAGATATTAAGACAGGACCAT	2760
Qy	2824	TATGAGTCTCTCAAGGAGAGAGATACCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTTTC	2883
Db	2761	TATGAGTCTCTCAAGGAGAGAGATACCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTTTC	2820
Qy	2884	CAAAATGGATTCGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGCTTTATTCGGGACCCAGGTC	2943
Db	2821	CAAAATGGATTCGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGCTTTATTCGGGACCCAGGTC	2880
Qy	2944	TTTGAAGACTATGACCTTCGCAAGAGCTGGTGGACTACATTTGACTGGAAGCCTTTCTTTGAT	3003
Db	2881	TTTGAAGACTATGACCTTCGCAAGAGCTGGTGGACTACATTTGACTGGAAGCCTTTCTTTGAT	2940
Qy	3004	GTCCTGGAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTAAACGACAAA	3063
Db	2941	GTCCTGGAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTAAACGACAAA	3000
Qy	3064	ACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCACAATATGCTGAACACACTG	3123
Db	3001	ACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCACAATATGCTGAACACACTG	3060
Qy	3124	ATTAGTCAAAAGAACTCCGGGCCCGGGTGTGGTTGGTTCCTGGCCAGCACAGAGTATC	3183
Db	3061	ATTAGTCAAAAGAACTCCGGGCCCGGGTGTGGTTGGTTCCTGGCCAGCACAGAGTATC	3120
Qy	3184	CAAGACGACATTTCACCTGTACGACAGGCTGTGTGCCCAAGCTGCAGAGCCCATAGCC	3243
Db	3121	CAAGACGACATTTCACCTGTACGACAGGCTGTGTGCCCAAGCTGCAGAGCCCATAGCC	3180
Qy	3244	ACTTTCATATGGTTAAAGCACAAGCTCAGAAAGACTCTGCCAGCACGAGCCATACTAC	3303
Db	3181	ACTTTCATATGGTTAAAGCACAAGCTCAGAAAGACTCTGCCAGCACGAGCCATACTAC	3240
Qy	3304	TGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACCTGGGCCCTGTTT	3363
Db	3241	TGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACCTGGGCCCTGTTT	3300
Qy	3364	GCCGTGTGCTGCTTTGGGGTAGAAGCTGAGCAAGGCCCTATGAGATATGATGTGACGAC	3423
Db	3301	GCCGTGTGCTGCTTTGGGGTAGAAGCTGAGCAAGGCCCTATGAGATATGATGTGACGAC	3360
Qy	3424	TACAGCAGCATCATGCTCAAGCGCTGGGGGACCGCTCGGCGAGAGCCCTTTGCAGAGAG	3483
Db	3361	TACAGCAGCATCATGCTCAAGCGCTGGGGGACCGCTCGGCGAGAGCCCTTTGCAGAGAG	3420
Qy	3484	CTCCATGAAAGAGTTTCGCCAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACGTC	3543
Db	3421	CTCCATGAAAGAGTTTCGCCAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACGTC	3480
Qy	3544	GCAGACTTCGGAAGTTTCGGGTACAGGCGCATCCGCCCGGCTCCTGGCTAACCCAGCCAG	3603
Db	3481	GCAGACTTCGGAAGTTTCGGGTACAGGCGCATCCGCCCGGCTCCTGGCTAACCCAGCCAG	3540
Qy	3604	CCCGACCAACCGAGAGCTCCACATGTGGAGCTCGCAGACATCCAGCAGTCTACAGGC	3663
Db	3541	CCCGACCAACCGAGAGCTCCACATGTGGAGCTCGCAGACATCCAGCAGTCTACAGGC	3600
Qy	3664	ATTAGGTTAAAGAAATCATTTAGCAATGGCACCTGCTTCAGCAGTCTCAGGCCCTACTTC	3723
Db	3601	ATTAGGTTAAAGAAATCATTTAGCAATGGCACCTGCTTCAGCAGTCTCAGGCCCTACTTC	3660
Qy	3724	TCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTTCCAAAGGATCAGGTTGAGAT	3783





Qy	1204	GCAGGATCAAGGAAGTTTGCTAARACTCATATGCGCAGGAACACTATGAAGAAGCCTTCTGT	1263	Qy	2284	AAGAGGCTGTGGCCACCTTATCCCTTTTCATGAAAAAAGAGAAGAAACACAGAGTG	2343
Db	1141	GCAGGATCAAGGAAGTTTGCTAARACTCATATGCGCAGGAACACTATGAAGAAGCCTTCTGT	1200	Db	2221	AAGAAGCTGTGGCCACCTTATCCCTTTTCATGAAAAAAGAGAAGAAACACAGAGTG	2280
Qy	1264	GTTCCCAAGTGCAGGTGGAAATGGAGACCCAGAGTGTGGATGTCAACATGGATGATGGC	1323	Qy	2344	CTTAAACGGCAGCTAGAGAAGAGGACCTTACACAGGCGCACCATCGTCTGGCCACTGTT	2403
Db	1201	GTTCCCAAGTGCAGGTGGAAATGGAGACCCAGAGTGTGGATGTCAACATGGATGATGGC	1260	Db	2281	CTTAAACGGCAGCTAGAGAAGAGGACCTTACACAGGCGCACCATCGTCTGGCCACTGTT	2340
Qy	1324	ATGCTAGATGGTCCAAAGTGCAATGACACAGATTTTGGCACTTAATTTGCTTCGGAGCCAGAC	1383	Qy	2404	AAAGGCGAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCGCAATAAT	2463
Db	1261	ATGCTAGATGGTCCAAAGTGCAATGACACAGATTTTGGCACTTAATTTGCTTCGGAGCCAGAC	1320	Db	2341	AAAGGCGAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCGCAATAAT	2400
Qy	1384	ATCCCAAGGTACCTTTGTGTCATCGACTCCTCCAAATTTTGTGTTGATTTGAAGCTGGGTTA	1443	Qy	2464	TTCCGAGTTAATGATTTAGGAGTCATGACTCCAATGTGATTAAGATACCTGAAAGCTGCTCTT	2523
Db	1321	ATCCCAAGGTACCTTTGTGTCATCGACTCCTCCAAATTTTGTGTTGATTTGAAGCTGGGTTA	1380	Db	2401	TTCCGAGTTAATGATTTAGGAGTCATGACTCCAATGTGATTAAGATACCTGAAAGCTGCTCTT	2460
Qy	1444	AAGTGCTGCCAAGGGAAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAGGAGAGGACGAC	1503	Qy	2524	GACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTTCCCTCCCTGGATGAATG	2583
Db	1381	AAGTGCTGCCAAGGGAAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAGGAGAGGACGAC	1440	Db	2461	GACCAAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTTCCCTCCCTGGATGAATG	2520
Qy	1504	TTCTTGGAGAAGGCCAGGAAGATTAAAGATATGGAGCTGCTATGTGGTGCATGGCTTTT	1563	Qy	2584	ATTTTGTTCGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATTGGAGGA	2643
Db	1441	TTCTTGGAGAAGGCCAGGAAGATTAAAGATATGGAGCTGCTATGTGGTGCATGGCTTTT	1500	Db	2521	ATTTTGTTCGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATTGGAGGA	2580
Qy	1564	GATGAAGAGACAGGCAACAGAACACAGACCAAAATCAGAGTGTGCACCCGGGCCCTAC	1623	Qy	2644	GCAACCACTTCAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGCACCTGTA	2703
Db	1501	GATGAAGAGACAGGCAACAGAACACAGACCAAAATCAGAGTGTGCACCCGGGCCCTAC	1560	Db	2581	GCAACCACTTCAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGCACCTGTA	2640
Qy	1624	CATCTGCTTGTGAAAAAATCGGCTTTAATCCAAATGACATATTTTGTGACCTTAATATC	1683	Qy	2704	ATCCATGCTCGACGGTCCAAAGAGTGTGGTGGTGTTCGCCAGCTGCCAGCTGTTAGATGAAT	2763
Db	1561	CATCTGCTTGTGAAAAAATCGGCTTTAATCCAAATGACATATTTTGTGACCTTAATATC	1620	Db	2641	ATCCATGCTCGACGGTCCAAAGAGTGTGGTGGTGTTCGCCAGCTGCCAGCTGTTAGATGAAT	2700
Qy	1684	CTAACCAATTGGGACTGGAAATGGAGAACACAACTTGTATGCAATTAATTTTATCCATGCA	1743	Qy	2764	CTAAAGGATGAATACTTTGAGGAAATCATCGAAGATATGAAGATATTAAGACGAGCCAT	2823
Db	1621	CTAACCAATTGGGACTGGAAATGGAGAACACAACTTGTATGCAATTAATTTTATCCATGCA	1680	Db	2701	CTAAAGGATGAATACTTTGAGGAAATCATCGAAGATATGAAGATATTAAGACGAGCCAT	2760
Qy	1744	ACAAAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAGGTCTTCCAACTTG	1803	Qy	2824	TATGAGTCTCTCAAGGAGAGGAGATCTTACCTTAAAGTCAAGCCAGAAAAAGTGTTC	2883
Db	1681	ACAAAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAGGTCTTCCAACTTG	1740	Db	2761	TATGAGTCTCTCAAGGAGAGGAGATCTTACCTTAAAGTCAAGCCAGAAAAAGTGTTC	2820
Qy	1804	TCCTTCTCTTCCGAGGAATGGAGCCATTCGAGAACATTCGATGGGGTTTCTTTTAC	1863	Qy	2884	CAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGGTC	2943
Db	1741	TCCTTCTCTTCCGAGGAATGGAGCCATTCGAGAACATTCGATGGGGTTTCTTTTAC	1800	Db	2821	CAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGGTC	2880
Qy	1864	CATGCAATCAAGTCTGGCATGGACATGGGATAGTGAATGCTGGAAAACTCCCTGTGTAT	1923	Qy	2944	TTTGAAGACTATGACTGCGAGAGCTGCTGAGCTTACATTTGACTGGAAGCCCTTTCTTTGAT	3003
Db	1801	CATGCAATCAAGTCTGGCATGGACATGGGATAGTGAATGCTGGAAAACTCCCTGTGTAT	1860	Db	2881	TTTGAAGACTATGACTGCGAGAGCTGCTGAGCTTACATTTGACTGGAAGCCCTTTCTTTGAT	2940
Qy	1924	GATGATATCCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCCCT	1983	Qy	3004	GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCCCAAGATATTTAAACGACAAA	3063
Db	1861	GATGATATCCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCCCT	1920	Db	2941	GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACAAA	3000
Qy	1984	GAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAGTC	2043	Qy	3064	ACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAAAATATGCTGGAACACACTG	3123
Db	1921	GAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAGTC	1980	Db	3001	ACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAAAATATGCTGGAACACACTG	3060
Qy	2044	ATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTGTG	2103	Qy	3124	ATTAAGTCAAAAGAAACTCCGGGCCCGGGGTGTGGTTCCTGCGCAGCACAGATATC	3183
Db	1981	ATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTGTG	2040	Db	3061	ATTAAGTCAAAAGAAACTCCGGGCCCGGGGTGTGGTTCCTGCGCAGCACAGATATC	3120
Qy	2104	AAGGCCATTGAAMAAACATATTTAGAGATCTGAGGAGCCAGGTTAAACCAAAAAAAA	2163	Qy	3184	CAAGACACATTCACCTGTGTACGAGAGCTGTGTGCCCGCCAGGCTGACAGGCCCATAGCC	3243
Db	2041	AAGGCCATTGAAMAAACATATTTAGAGATCTGAGGAGCCAGGTTAAACCAAAAAAAA	2100	Db	3121	CAAGACACATTCACCTGTGTACGAGAGCTGTGTGCCCGCCAGGCTGACAGGCCCATAGCC	3180
Qy	2164	TATCCCCGACCTCTCAATATAATTAAGAGCCCTCTGATGGAATGAAATTTGCTG	2223	Qy	3244	ACTTTCTATGGGTTAAGGCAACAGGCTGAGAAAGCACTCTGCGACGACGGAGCCCATATAC	3303
Db	2101	TATCCCCGACCTCTCAATATAATTAAGAGCCCTCTGATGGAATGAAATTTGCTG	2160	Db	3181	ACCTTCTATGGGTTAAGGCAACAGGCTGAGAAAGACTCTGCGACGACGGAGCCCATATAC	3240
Qy	2224	GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCGGGTTATG	2283	Qy	3304	TGCTCTCAGACTTTCATCGCTCCCTTGGCATCTCGCTACCTGCTGCTGGGCTGTTT	3363
Db	2161	GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCGGGTTATG	2220	Db	3241	TGCTCTCAGACTTTCATCGCTCCCTTGGCATCTCGCTACCTGCTGCTGGGCTGTTT	3300
				Qy	3364	GCCGTTGCTCTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGCTGACGAC	3423

```
|||||
Db 3301 GCCGTTGCCCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCCTATGAGGATGATGGTGACGAC 3360
Qy 3424 TACAGCAGCATCATGTGTCAAGCGCTCGGGGACCCGGCTGGCAGAGGCTTTGCAGAAAGAG 3483
Db 3361 TACAGCAGCATCATGTGTCAAGCGCTCGGGGACCCGGCTGGCAGAGGCTTTGCAGAAAGAG 3420
Qy 3484 CTCCATGAAGAGTTCCCGGAGAACTGTGGGCTTACTGTGGCAGTGAAGAGCTGGACGTC 3543
Db 3421 CTCCATGAAGAGTTCCCGGAGAACTGTGGGCTTACTGTGGCAGTGAAGAGCTGGACGTC 3480
Qy 3544 GCAGACCTGGGAAGTTGCGGTACAAGGGCATCCGCCGGCTCCTGGCTACCCCGCCAG 3603
Db 3481 GCAGACCTGGCGAGCTGCGGTACAAGGGCATCCGCCGGCTCCTGGCTACCCCGCCAG 3540
Qy 3604 CCCGACCACACCCGAGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC 3663
Db 3541 CCCGACCACACCCGAGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC 3600
Qy 3664 ATTAGGTTAACAGAAATCATTTAGCAATGGCACCTGCTTCAGGAGTCTCAGGCTCTACTTC 3723
Db 3601 ATTAGGTTAACAGAAATCATTTAGCAATGGCACCTGCTTCAGGAGTCTCAGGCTCTACTTC 3660
Qy 3724 TCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAGGATCAGGTTGAGGAT 3783
Db 3661 TCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAGGATCAGGTTGAGGAT 3720
Qy 3784 TATGCATTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTTTG 3843
Db 3721 TATGCATTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTTTG 3780
Qy 3844 GGATATGATACAGACTAA 3861
Db 3781 GGATATGATACAGACTAA 3798
```

Search completed: March 6, 2006, 20:35:18  
Job time.: 19274 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 14:47:47 ; Search time 1474 Seconds  
(without alignments)  
17719.763 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcacctgtggagacgacg.....ctcaaggaaatacaacctag 3919

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_21.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2002bs.\*
- 8: Geneseq2003as.\*
- 9: Geneseq2003bs.\*
- 10: Geneseq2003cs.\*
- 11: Geneseq2003ds.\*
- 12: Geneseq2004as.\*
- 13: Geneseq2004bs.\*
- 14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3884.8	99.1	7122	4	AAF31100
2	3884.8	99.1	7122	5	AAF31100 Methionin
3	3884.8	99.1	7122	14	AAC91207 Human met
4	3883.2	99.1	7122	14	Aea63658 Methionin
5	3793.8	96.8	3920	2	AAF31101 Methionin
6	3785.2	96.6	3798	10	Aav34063 Human met
7	926.8	23.6	3780	10	ADI39109 Homo sapi
8	914.6	23.3	3786	8	ADI39113 Agrobacte
9	911.8	23.3	3684	8	Aca37652 Prokaryot
10	903	23.0	3762	11	Abd14844 Pseudomon
11	903	23.0	3795	11	Abd15089 Pseudomon
12	902.6	23.0	3669	8	ACA44722 Prokaryot
13	901.4	23.0	3705	8	ACA42279 Prokaryot
14	901.4	23.0	3705	10	ADI39093 Pseudomon
15	901	23.0	3750	10	ADFO2792 Bacterial
16	899	22.9	3681	8	ACA32117 Prokaryot
17	895.8	22.9	3702	9	ADA30291 DNA encod
18	894	22.8	3705	8	ACA44248 Prokaryot
19	892.2	22.8	3717	8	ACA45414 Prokaryot

20	883.8	22.6	3684	8	ACA32726	Aca32726 Prokaryot
21	883.8	22.6	3684	10	ADI39086	Adi39086 Escherich
22	883.2	22.5	3681	8	ACA52977	Aca52977 Prokaryot
23	883.2	22.5	3681	10	ADI39082	Adi39082 Vibrio ch
24	877.6	22.4	3645	10	ADI39117	Adi39117 Chlorobiu
25	877.6	22.4	3645	10	ADI39099	Adi39099 Chlorobiu
26	873.2	22.3	3711	10	ADI39091	Adi39091 Pseudomon
27	867.8	22.1	3714	10	ADI39095	Adi39095 Nitrosomo
28	867.2	22.1	3684	8	ACA51700	Aca51700 Prokaryot
29	867	22.1	3822	10	ADI39084	Adi39084 Sinorhizo
30	865.2	22.1	3681	10	ADI39111	Adi39111 Vibrio fi
31	859	21.9	3681	8	ACA36141	Aca36141 Prokaryot
32	857.6	21.9	3771	8	ACA51042	Adi39088 Salmonell
33	857.6	21.9	3771	10	ADI39088	Adi39088 Prokaryot
34	854.4	21.8	3696	8	ACA53745	Aca53745 Prokaryot
35	840.8	21.5	3759	10	ADI39107	Adi39107 Rhodobact
36	835	21.3	3642	10	ADI39103	Adi39103 Clostridi
37	763.8	19.5	3777	10	ADI39101	Adi39101 Deinococc
38	749.2	19.1	3774	10	ADI39097	Adi39097 Bordetell
39	745.6	19.0	2796	8	ACA24643	Aca24643 Prokaryot
40	708.6	18.1	3157	8	ACA48896	Aca48896 Prokaryot
41	701	17.9	3138	11	ACH95623	Ach95623 Klebsiell
42	650.4	16.6	3954	10	ADI39105	Adi39105 Caulobact
43	621.4	15.9	2718	10	ADI39115	Adi39115 Ralstonia
44	615.8	15.7	2349	11	ABD15289	Abd15289 Pseudomon
45	477.8	12.2	1779	8	ACA26885	Aca26885 Prokaryot

#### ALIGNMENTS

RESULT 1  
AAF31100  
ID AAF31100 standard; CDNA; 7122 BP.  
XX  
AC AAF31100;  
XX  
DT 27-APR-2001 (first entry)  
XX  
DE Methionine synthase coding sequence #1.  
XX  
KW Analyte-binding enzyme; analyte analysis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200102600-A2.  
XX  
PD 11-JAN-2001.  
XX  
PF 30-JUN-2000; 2000WO-US018057.  
XX  
PR 06-JUL-1999; 99US-00347878.  
PR 06-DEC-1999; 99US-00457205.  
XX  
(GEAT ) GEN ATOMICS.  
PA  
XX  
PI Yuan C;  
XX  
DR WPI; 2001-071583/08.  
XX  
PT Assaying method, useful for prognosis and diagnosis of disease, comprises  
PT contacting sample with a mutant analyte-binding enzyme and detecting  
PT binding.  
XX  
PS Disclosure; Page; 187pp; English.  
XX  
CC The present invention relates to a method for assaying an analyte in a  
CC sample comprising: contacting the sample with a mutant analyte-binding  
CC enzyme which has binding affinity for the analyte or an immediate analyte  
CC enzymatic conversion product but has attenuated catalytic activity; and  
CC detecting resulting binding. The method is useful in monitoring  
CC biological systems/processes, or prognosis/diagnosis of disease caused by  
CC imbalances of the analytes. The present sequence is a coding sequence

CC used in the present invention. Note: the present sequence is not shown in  
CC the specification, but was from Genbank, using information given in the  
CC specification

XX Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;

Query Match 99.1%; Score 3884.8; DB 4; Length 7122;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3904; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

Qy	2	GTCACTGTGGAGACGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTGCACA	61
Db	225	GTCACTGTGGAGACGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTGCACA	284
Qy	62	ACATGTCAACCGCGCTCCAGACCTGTGCGAACCGGAGGCTCTGAAGAAACCTGCGGG	121
Db	285	ACATGTCAACCGCGCTCCAGACCTGTGCGAACCGGAGGCTCTGAAGAAACCTGCGGG	344
Qy	122	ATGAGATCAATGCCATTTCTGAGAGAGGATTATGGTGTGATGGAGGGATGGGACCA	181
Db	345	ATGAGATCAATGCCATTTCTGAGAGAGGATTATGGTGTGATGGAGGGATGGGACCA	404
Qy	182	TGATCCAGCGGAGAAAGCTAAACGAGAACCTTCCGAGGTCAGGAATTTAAAGATCATG	241
Db	405	TGATCCAGCGGAGAAAGCTAAACGAGAACCTTCCGAGGTCAGGAATTTAAAGATCATG	464
Qy	242	CCAGGCGCTGAAAGGCAAGATGACATTTTAAGTATAACTCAGCCTGATGTCATTTACC	301
Db	465	CCAGGCGCTGAAAGGCAAGATGACATTTTAAGTATAACTCAGCCTGATGTCATTTACC	524
Qy	302	AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA	361
Db	525	AAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA	584
Qy	362	GCACCTAGTATTGCCAAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT	421
Db	585	GCACCTAGTATTGCCAAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT	644
Qy	422	GCTCTGAGGAGTGGCCGAGAAAGGTGCGGAGGAGGTAACCTCTCAGACAGGAATTAAGA	481
Db	645	GCTCTGAGGAGTGGCCGAGAAAGGTGCGGAGGAGGTAACCTCTCAGACAGGAATTAAGA	704
Qy	482	GGTTTGTGGCAGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG	541
Db	705	GGTTTGTGGCAGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG	764
Qy	542	AAAGCGCGGATTAAGGAACATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAGG	601
Db	765	AAAGCGCGGATTAAGGAACATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAGG	824
Qy	602	CCTAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTACTG	661
Db	825	CCTAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTACTG	884
Qy	662	CCATGCAAGCAGCCTGTGTGCACTCCAAATCTTTTGGAGAGAAATATGTCCTCCC	721
Db	885	CCATGCAAGCAGCCTGTGTGCACTCCAAATCTTTTGGAGAGAAATATGTCCTCCC	944
Qy	722	GGCCTATCTTTATTTTCAGGACGATCGTTGATAAAGTGGGCGGACTCTTTCCGACAGA	781
Db	945	GGCCTATCTTTATTTTCAGGACGATCGTTGATAAAGTGGGCGGACTCTTTCCGACAGA	1004
Qy	782	CAGGAGGGGATTTGTATCAGCGGTGTCTCATGGAGAACCACTCTGCAATTTGGATTAATTT	841
Db	1005	CAGGAGGGGATTTGTATCAGCGGTGTCTCATGGAGAACCACTCTPACATTTGGATTAATTT	1064
Qy	842	GTGCTTTGGGTGCAGCTCAGATGAGACCTTTTATTTGAATTAATTTGGAATTTGACACAG	901
Db	1065	GTGCTTTGGGTGCAGCTGAAATGAGACCTTTTATTTGAATTAATTTGGAATTTGACACAG	1124
Qy	902	CCTATGCTCTGTGTTATCCCAATGAGGCTTTTCCCAACACCTTTTGGTGACTATGATGAAA	961
Db	1125	CCTATGCTCTGTGTTATCCCAATGAGGCTTTTCCCAACACCTTTTGGTGACTATGATGAAA	1184

Qy	962	CGCTTCTATGATGGCCAAAGCACTAAAGGATTTTGTATGGATGGCTTGGTCAATATAG	1021
Db	1185	CGCTTCTATGATGGCCAAAGCACTAAAGGATTTTGTATGGATGGCTTGGTCAATATAG	1244
Qy	1022	TTGGAGGATGCTGTGGTCAACACCCAGATCATATCAGGGGAAATTTGCTGAAGCTGTGAAA	1081
Db	1245	TTGGAGGATGCTGTGGTCAACACCCAGATCATATCAGGGGAAATTTGCTGAAGCTGTGAAA	1304
Qy	1082	ATTGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGGTC	1141
Db	1305	ATTGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGGTC	1364
Qy	1142	TAGAGCCCTTCAGATTTGGACCGTACACCACTTTGTTTAACTATGGAGAGCGCTGTAATG	1201
Db	1365	TAGAGCCCTTCAGATTTGGACCGTACACCACTTTGTTTAACTATGGAGAGCGCTGTAATG	1424
Qy	1202	TTGCAAGGATCAAGGAAAGTTTGTCTAAACTCATCATGCGAGGAACTATGAAGAGCCTTGT	1261
Db	1425	TTGCAAGGATCAAGGAAAGTTTGTCTAAACTCATCATGCGAGGAACTATGAAGAGCCTTGT	1484
Qy	1262	GTGTTGCCAAAGTGCAGGTGGAATGGGAGCCAGGTGTTGGATGTCAAACATGGATGATG	1321
Db	1485	GTGTTGCCAAAGTGCAGGTGGAATGGGAGCCAGGTGTTGGATGTCAAACATGGATGATG	1544
Qy	1322	GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTTCGAACTTAAATTTGCTTCCGAGCCAG	1381
Db	1545	GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTTCGAACTTAAATTTGCTTCCGAGCCAG	1604
Qy	1382	ACATCGCAAAAGTACCTTTTGTGATCGACTCTCTCCAAATTTTGTGTAATGAAGCTGGGT	1441
Db	1605	ACATCGCAAAAGTACCTTTTGTGATCGACTCTCTCCAAATTTTGTGTAATGAAGCTGGGT	1664
Qy	1442	TAAAGTCTGCTCCAAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAAGGAGAGCG	1501
Db	1665	TAAAGTCTGCTCCAAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAAGGAGAGCG	1724
Qy	1502	ACTTCTTGGAGAGAGGCGCAGGAAGATTAAGAAAGTATGGAGCTGCTATGGTGTCTATGGCTT	1561
Db	1725	ACTTCTTGGAGAGAGGCGCAGGAAGATTAAGAAAGTATGGAGCTGCTATGGTGTCTATGGCTT	1784
Qy	1562	TTGATGAAGAGGAGCAGGCAACAGAAAACAGACACAGAAAATCAGAGTGTGACCCGGGCT	1621
Db	1785	TTGATGAAGAGGAGCAGGCAACAGAAAACAGACACAGAAAATCAGAGTGTGACCCGGGCT	1844
Qy	1622	ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGACCTTAATA	1681
Db	1845	ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA	1904
Qy	1682	TCCTAACCATTTGGGACTGGGAATGGAGAACACAACTTTGTATGCCATTAATTTTATCCATG	1741
Db	1905	TCCTAACCATTTGGGACTGGGAATGGAGAACACAACTTTGTATGCCATTAATTTTATCCATG	1964
Qy	1742	CAACAAAGTCAATTAAGAAACATTTACTGAGGACAGATAAGTGGAGGCTTTTCCAACT	1801
Db	1965	CAACAAAGTCAATTAAGAAACATTTACTGAGGACAGATAAGTGGAGGCTTTTCCAACT	2024
Qy	1802	TGTCCTTCTCTCCGAGGAAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTCCTTTT	1861
Db	2025	TGTCCTTCTCTCCGAGGAAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTCCTTTT	2084
Qy	1862	ACCATGCAATCAAGTCTGGCATGGACATGGAGTAGTGAAATGCTGGGAAACCTCCTCTGT	1921
Db	2085	ACCATGCAATCAAGTCTGGCATGGACATGGGATAGTGAAATGCTGGGAAACCTCCTCTGT	2144
Qy	1922	ATGATGATATCCATAAGGAACCTTCTGCAAGCTCTCTGGAAGATCTCATCTCGGAATAAGACC	1981
Db	2145	ATGATGATATCCATAAGGAACCTTCTGCAAGCTCTCTGGAAGATCTCATCTCGGAATAAGACC	2204
Qy	1982	CTGAGGCCCATGAGAGCTCTTATGTTATGCCAGACTCAAGGACACAGGAGGAGGAAG	2041
Db	2205	CTGAGGCCCATGAGAGCTCTTATGTTATGCCAGACTCAAGGACACAGGAGGAGGAAG	2264

QY 2042 TCATTGAGCTGATGAGTGGAGAAATGCGCCCTGTGGAAGACGCTTGAGTATGCCCTTG 2101  
Db 2265 TCATTGAGCTGATGAGTGGAGAAATGCGCCCTGTGGAAGACGCTTGAGTATGCCCTTG 2324  
QY 2102 TGAAGGCGATTGAAAAACATATTATTCAGGATACCTGAGGAAAGCCAGGTAAACCAAAAAA 2161  
Db 2325 TGAAGGCGATTGAAAAACATATTATTCAGGATACCTGAGGAAAGCCAGGTAAACCAAAAAA 2384  
QY 2162 AATATCCCGACCTCTCAATATATTAAGAGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
Db 2385 AATATCCCGACCTCTCAATATATTAAGAGACCCCTGATGAATGGAATGAAATTTGTTG 2444  
QY 2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTATTAAGTCAAGCCGGGTGA 2281  
Db 2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTATTAAGTCAAGCCGGGTGA 2504  
QY 2282 TGAAGAAAGGCTGTTGGCCACCTTATCCCTTTTATGGAAGAAAGAAAGAAACACAGAG 2341  
Db 2505 TGAAGAAAGGCTGTTGGCCACCTTATCCCTTTTATGGAAGAAAGAAAGAAACACAGAG 2564  
QY 2342 TGCTTAACCGGCACAGTAGAAGAGAGACCCCTTACACAGGCGCACCATCGTCTGCCCACTG 2401  
Db 2565 TGCTTAACCGGCACAGTAGAAGAGAGAGACCCCTTACACAGGCGCACCATCGTCTGCCCACTG 2624  
QY 2402 TTAAGGCGGACGTCACGACATAGGCAAGACATAGTGGAGTAGTCTTGGCTGCAATA 2461  
Db 2625 TTAAGGCGGACGTCACGACATAGGCAAGACATAGTGGAGTAGTCTTGGCTGCAATA 2684  
QY 2462 ATTTCCGAGTTATTGATTTAGGAGTCATGACTCCTCATGTGATAGATACGAAAGCTGCTC 2521  
Db 2685 ATTTCCGAGTTATTGATTTAGGAGTCATGACTCCTCATGTGATAGATACGAAAGCTGCTC 2744  
QY 2522 TTGACCAAAAGCAGATATTAATTTGGCCTGTGAGGACTCATCACTTCCCTCCCTGAGTAA 2581  
Db 2745 TTGACCAAAAGCAGATATTAATTTGGCCTGTGAGGACTCATCACTTCCCTCCCTGAGTAA 2804  
QY 2582 TGATTTTGTGCAAGGAAATGAGAGATAGTACTAAGATTCATAGTGTGATGAGGAG 2641  
Db 2805 TGATTTTGTGCAAGGAAATGAGAGATAGTACTAAGATTCATAGTGTGATGAGGAG 2864  
QY 2642 GAGCAACCACTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACACTG 2701  
Db 2865 GAGCAACCACTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACACTG 2924  
QY 2702 TAATCCATGTCTCGAGCGCTGCAAGAGTGTGGTGTGTTCCAGCTGTAGATGAAA 2761  
Db 2925 TAATCCATGTCTCGAGCGCTGCAAGAGTGTGGTGTGTTCCAGCTGTAGATGAAA 2984  
QY 2762 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATATGAAGATTTAGACAGACC 2821  
Db 2985 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATATGAAGATTTAGACAGACC 3044  
QY 2822 ATTTAGTCTCTCAAGGAGGAGATACCTTACCTTTAGCTCAAGCCAGAAAGTGGTT 2881  
Db 3045 ATTTAGTCTCTCAAGGAGGAGATACCTTACCTTTAGCTCAAGCCAGAAAGTGGTT 3104  
QY 2882 TCCAAATGAGTGTGCTGTAACCTCACAGTGAAGCCACCGTTTATTTGGGACCCAGG 2941  
Db 3105 TCCAAATGAGTGTGCTGTAACCTCACAGTGAAGCCACCGTTTATTTGGGACCCAGG 3164  
QY 2942 TCTTTGAAGACTATGACTGACAGAGCTGTGGTGAATCATGACTGGAAGCTTTCTTTG 3001  
Db 3165 TCTTTGAAGACTATGACTGACAGAGCTGTGGTGAATCATGACTGGAAGCTTTCTTTG 3224  
QY 3002 ATGCTGCGCAGCTCGGGGCAAGTACCGGAATCAGGCTTCCCAAGATTTTAACGACA 3061  
Db 3225 ATGCTGCGCAGCTCGGGGCAAGTACCGGAATCAGGCTTCCCAAGATTTTAACGACA 3284  
QY 3062 AAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAATATGCTGAACACAC 3121  
Db 3285 AAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAATATGCTGAACACAC 3344  
QY 3122 TGATTAGTCAAAAGAAATCCGGGCCCCGGGGTGTGGTGTCTGCGCCAGCAGAGTA 3181

Db 3345 TGATTAGTCAAAAGAAATCTCCGGCCCCGGGTGTGGTGTCTGCGCCAGCAGAGTA 3404  
QY 3182 TCCAAAGACGACATTTCACTGTAGCGAGAGGCTGTCTGTGCCCCAGGCTGCGAGGCCCATAG 3241  
Db 3405 TCCAAAGACGACATTTCACTGTAGCGAGAGGCTGTCTGTGCCCCAGGCTGCGAGGCCCATAG 3464  
QY 3242 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3301  
Db 3465 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3524  
QY 3302 ACTGCGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGCGATCCGCTACTACTCGGGCTGT 3361  
Db 3525 ACTGCGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGCGATCCGCTACTACTCGGGCTGT 3584  
QY 3362 TTGCCCTTGCCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3421  
Db 3585 TTGCCCTTGCCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3644  
QY 3422 ACTACAGCAGCATCATGCTCAGGCGCTGGGGACCGGCTGGCAGAGGCCCTTTGCGAAG 3481  
Db 3645 ACTACAGCAGCATCATGCTCAGGCGCTGGGGACCGGCTGGCAGAGGCCCTTTGCGAAG 3704  
QY 3482 AGCTCCATGAAGAGTTCGCGGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGACG 3541  
Db 3705 AGCTCCATGAAGAGTTCGCGGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGACG 3764  
QY 3542 TCGCAGACCTGCGAAGGTTGCGGTACAAGGCAATCCGCGCGGCTCCTGGCTACCCAGCC 3601  
Db 3765 TCGCAGACCTGCGCAGGCTGCGGTACAAGGCAATCCGCGCGGCTCCTGGCTACCCAGCC 3824  
QY 3602 AGCCGACACACACCGAAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3661  
Db 3825 AGCCGACACACACCGAAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3884  
QY 3662 GCATTAGGTTACAGAAATCAATTAGCAATGCGACCTGCTTCCAGCAGTCTCAGGCTCTACT 3721  
Db 3885 GCATTAGGTTACAGAAATCAATTAGCAATGCGACCTGCTTCCAGCAGTCTCAGGCTCTACT 3944  
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 3781  
Db 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 4004  
QY 3782 ATTATGCAATGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 3841  
Db 4005 ATTATGCAATGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 4064  
QY 3842 TGGGATATGATACAGACTAACTTTTGTGTGGCTTTTGTGTGGCTTTTGTGTGTGATGATCCT 3901  
Db 4065 TGGGATATGATACAGACTAACTTTTGTGTGGCTTTTGTGTGGCTTTTGTGTGTGATGATCCT 4122  
QY 3902 CAAGGAAATACAACTTAG 3919  
Db 4123 CAAGGAAATACAACTTAG 4140

RESULT 2  
AAC91207  
ID AAC91207 standard; DNA; 7122 BP.  
XX AC AAC91207;  
XX AC AAC91207;  
DT 20-MAR-2001 (first entry)  
XX Human methionine synthase gene SEQ ID NO: 2.  
DE Human; schizophrenia; developmental disorder; spina bifida cystica;  
XX Tourette's syndrome; bipolar illness; autism; conduct disorder;  
KW attention deficit disorder; obsessive compulsive disorder;  
KW chronic multiple tic syndrome; learning disorder; polymorphism; da.  
XX Homo sapiens.  
OS XX

PN WO200071754-A1.  
 XX 30-NOV-2000.  
 XX 24-MAY-2000; 2000WO-US014354.  
 XX 25-MAY-1999; 99US-00318448.  
 XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
 XX Johnson WG, Stenroos ES;  
 XX WPI; 2001-025174/03.  
 XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
 PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
 PT and environmental variables affecting an individual and then comparing  
 PT these DS with reference DS.  
 XX  
 XX Disclosure; Page 115-117; 156pp; English.  
 XX  
 CC The present invention provides a novel method of estimating the  
 CC susceptibility of an individual to a developmental disorder using genetic  
 CC and environmental variables. The method can be used in the diagnosis,  
 CC prevention and treatment of disorders such as schizophrenia, spina bifida  
 CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
 CC attention deficit disorder, obsessive compulsive disorder, chronic  
 CC multiple tic syndrome and learning disorders such as dyslexia  
 XX  
 XX Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;  
 SQ  
 Query Match 99.1%; Score 3884.8; DB 5; Length 7122;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 3504; Conservative 0; Mismatches 12; Indels 2; Gaps 1;  
 QY 2 GTCACCTGTGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61  
 DB 225 GTCACCTGTGAGAGCAGCTCTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 284  
 QY 62 ACATGTACCCGCGCTCCAGACCTGTGCGAACCCGAGAGGTCTGAAGAAACCCCTGGGG 121  
 DB 285 ACATGTACCCGCGCTCCAGACCTGTGCGAACCCGAGAGGTCTGAAGAAACCCCTGGGG 344  
 QY 122 ATGAGATCAATGCCATTCTGCAGAGAGAGATTATGCTGCTGATGAGGAGTGGGACCA 181  
 DB 345 ATGAGATCAATGCCATTCTGCAGAGAGAGATTATGCTGCTGATGAGGAGTGGGACCA 404  
 QY 182 TGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241  
 DB 405 TGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 464  
 QY 242 CCAGCGCGCTGAAGGCAACATGACATTTAAAGTATAACTCAGGCTGATGTCAATTACC 301  
 DB 465 CCAGCGCGCTGAAGGCAACATGACATTTAAAGTATAACTCAGGCTGATGTCAATTACC 524  
 QY 302 AAATCCATAAGGATACCTTGTGCTGGGCGGAGATATCATTTGAACAATACTTTTAGCA 361  
 DB 524 AAATCCATAAGGATACCTTGTGCTGGGCGGAGATATCATTTGAACAATACTTTTAGCA 584  
 QY 362 GCACCTAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 421  
 DB 585 GCACCTAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 644  
 QY 422 GCTCTGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACCTCTCCAGACAGGAATTAAGA 481  
 DB 645 GCTCTGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACCTCTCCAGACAGGAATTAAGA 704  
 QY 482 GGTTCGCGAGGGGCTCTGGGTCGACTTAATAGACACTCTCTGTGTCCCATCTGTGG 541  
 DB 705 GGTTCGCGAGGGGCTCTGGGTCGACTTAATAGACACTCTCTGTGTCCCATCTGTGG 764  
 QY 542 AAAGCCCGGATTTATAGGAACATCACTTTGATGAGCTTTGTTGAACATACCAAGACGAGG 601

DB 765 AAAGCCCGGATTTATAGGAACATCACATTTGATGAGCTTTGTTGAAGCATATACAGAGCAGG 824  
 QY 602 CCAAAGGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGATAAACTATTTTGTGATCTG 661  
 DB 825 CCAAAGGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGATAAACTATTTTGTGATCTG 884  
 QY 662 CCAATGCCAAGGAGCGCTTGTGCTCACTCCAAATCTTTTGTGAGGAGAAATATGCTCCCC 721  
 DB 885 CCAATGCCAAGGAGCGCTTGTGCTCACTCCAAATCTTTTGTGAGGAGAAATATGCTCCCC 944  
 QY 722 GGCTATCTTTATTTTTCAGGACGATCGTTGATATAAGTGGCGGACTCTTTCCGACAGA 781  
 DB 945 GGCTATCTTTATTTTCAGGACGATCGTTGATATAAGTGGCGGACTCTTTCCGACAGA 1004  
 QY 782 CAGGAGAGGATTTGTCTCATCAGCGTGTCTCATGAGAAACCACTCTGCTATTTGATTAATTT 841  
 DB 1005 CAGGAGAGGATTTGTCTCATCAGCGTGTCTCATGAGAAACCACTCTACATTTGGATTAATTT 1064  
 QY 842 GTGCTTTGGGTGCGAGCTGAGATGAGACCTTTTATTTGAAATTAATTTGGAATTTGTAACAAG 901  
 DB 1065 GTGCTTTGGGTGCGAGCTGAAATGAGACCTTTTATTTGAAATTAATTTGGAATTTGTAACAAG 1124  
 QY 902 CCTATGCTCTGTATATCCCAATGCGAGTCTTCCCAACACCTTTGGTGAATGATGAAA 961  
 DB 1125 CCTATGCTCTGTATATCCCAATGCGAGTCTTCCCAACACCTTTGGTGAATGATGAAA 1184  
 QY 962 CGCCTTCTATGATGGCCCAACCACTTAAAGATTTTGTCTATGAGTGGCTCTGCTCAATATAG 1021  
 DB 1185 CGCCTTCTATGATGGCCCAACCACTTAAAGATTTTGTCTATGAGTGGCTCTGCTCAATATAG 1244  
 QY 1022 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1081  
 DB 1245 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1304  
 QY 1082 ATTGTAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141  
 DB 1305 ATTGTAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1364  
 QY 1142 TAGAGCCCTTCAGATTTGGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAAATG 1201  
 DB 1365 TAGAGCCCTTCAGATTTGGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAAATG 1424  
 QY 1202 TTGCAAGATCAAGGAAGTTTGTCTAACTCATCTGCGAGGAAACTATGAAGAACCTTTGT 1261  
 DB 1425 TTGCAAGATCAAGGAAGTTTGTCTAACTCATCTGCGAGGAAACTATGAAGAACCTTTGT 1484  
 QY 1262 GTGTTCGCAAGTGCAGGTGGAATGGAGCCCAAGTGTGGATGTCAACATGATGATG 1321  
 DB 1485 GTGTTCGCAAGTGCAGGTGGAATGGAGCCCAAGTGTGGATGTCAACATGATGATG 1544  
 QY 1322 GCATGCTAGATGGTCCCAAGTGCATGACCAAGTTTTCGCACTTAATTTGCTTCCGAGCCAG 1381  
 DB 1545 GCATGCTAGATGGTCCCAAGTGCATGACCAAGTTTTCGCACTTAATTTGCTTCCGAGCCAG 1604  
 QY 1382 ACATCGCAAGGTAACCTTTGTGCACTCGACTCTCCAATTTTGTCTGATTTGAAGCTGGGT 1441  
 DB 1605 ACATCGCAAGGTAACCTTTGTGCACTCGACTCTCCAATTTTGTCTGATTTGAAGCTGGGT 1664  
 QY 1442 TAAAGTGTGCAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGACG 1501  
 DB 1665 TAAAGTGTGCAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGACG 1724  
 QY 1502 ACTTCTTGGAGAGGCGGAGGAGATTTAAAGTATGAGCTGCTATGGTGTCTAGGCTT 1561  
 DB 1725 ACTTCTTGGAGAGGCGGAGGAGATTTAAAGTATGAGCTGCTATGGTGTCTAGGCTT 1784  
 QY 1562 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATTCAGAGTGTGCAACCCGGGCT 1621  
 DB 1785 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATTCAGAGTGTGCAACCCGGGCT 1844  
 QY 1622 ACCATCTGCTTTGTGAAAAAACTGGGCTTTAATCAAAATGACATTTATTTTGAACCTTAATA 1681  
 DB 1845 ACCATCTGCTTTGTGAAAAAACTGGGCTTTAATCAAAATGACATTTATTTTGAACCTTAATA 1904

Qy	1682	TCCTAACCACTTGGGACTGGAAATGAGGAAACACAACCTGTGTATGCCATTAAATTTTATCCATG	1741
Db	1905	TCCTAACCACTTGGGACTGGAAATGAGGAAACACAACCTGTGTATGCCATTAAATTTTATCCATG	1964
Qy	1742	CAACAAAAGTCAATTAAGAAAACATTAACCTGGAGCCAGAAATAGTGGAGGTCCTTCCAACT	1801
Db	1965	CAACAAAAGTCAATTAAGAAAACATTAACCTGGAGCCAGAAATAGTGGAGGTCCTTCCAACT	2024
Qy	1802	TGCTCTCTCTCCGAGGAATGGAAGCCATTCGAGAAAGCAATGCATGGGGTTTTCCTTT	1861
Db	2025	TGCTCTCTCTCCGAGGAATGGAAGCCATTCGAGAAAGCAATGCATGGGGTTTTCCTTT	2084
Qy	1862	ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT	1921
Db	2085	ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT	2144
Qy	1922	ATGATGATATCCATAAGGAACCTTCGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC	1981
Db	2145	ATGATGATATCCATAAGGAACCTTCGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC	2204
Qy	1982	CTGAGGCCACTGAGAAAGTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG	2041
Db	2205	CTGAGGCCACTGAGAAAGTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG	2264
Qy	2042	TCATTCAGACTGATGAGGAAATGAGCCCTGTGCAAGAAAGCGCTTGAGTATGCCCTTG	2101
Db	2265	TCATTCAGACTGATGAGGAAATGAGCCCTGTGCAAGAAAGCGCTTGAGTATGCCCTTG	2324
Qy	2102	TGAAGGCAATGAAAACATATTTAGGATACCTGAGGAAAGCCAGGTTAAACCAAAAA	2161
Db	2325	TGAAGGCAATGAAAACATATTTAGGATACCTGAGGAAAGCCAGGTTAAACCAAAAA	2384
Qy	2162	AATATCCCGGACCTCTCAATATTAATGAAAGACCCCTGATCAATGGAATGAAATTTGTTG	2221
Db	2385	AATATCCCGGACCTCTCAATATTAATGAAAGACCCCTGATCAATGGAATGAAATTTGTTG	2444
Qy	2222	GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGGTTA	2281
Db	2445	GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGGTTA	2504
Qy	2282	TGAAGAGGCTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGGAAAGACCAAGAG	2341
Db	2505	TGAAGAGGCTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGGAAAGACCAAGAG	2564
Qy	2342	TGCTTAAAGGACAGTGAAGAAAGAGGACCTTACCAGGCAACCATCTGTGCTGGCCACTG	2401
Db	2565	TGCTTAAAGGACAGTGAAGAAAGAGGACCTTACCAGGCAACCATCTGTGCTGGCCACTG	2624
Qy	2402	TAAAGGCGAGTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGGCTGCAATA	2461
Db	2625	TAAAGGCGAGTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGGCTGCAATA	2684
Qy	2462	ATTTCCGAGTATTTGATTTAGAGTCACTGCTCAATGATGAAGATGCTGAAAGCTGCTC	2521
Db	2685	ATTTCCGAGTATTTGATTTAGAGTCACTGCTCAATGATGAAGATGCTGAAAGCTGCTC	2744
Qy	2522	TGACCAACAAGCAGATATAATGSCCTGTGAGGACTCATCACTCTTCCCTGGATGAAA	2581
Db	2745	TGACCAACAAGCAGATATAATGSCCTGTGAGGACTCATCACTCTTCCCTGGATGAAA	2804
Qy	2582	TGATTTTGTGTCGAAGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATGGAG	2641
Db	2805	TGATTTTGTGTCGAAGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATGGAG	2864
Qy	2642	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2865	GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	2924
Qy	2702	TAATCCATGCTGAGCGGCTCAAGAGTGTGGTGTGTTCCAGCTGTAGTAGAAA	2761
Db	2925	TAATCCATGCTGAGCGGCTCAAGAGTGTGGTGTGTTCCAGCTGTAGTAGAAA	2984

Qy	2762	ATCTAAGAGTGAATATCTTTGAGGAATCATGGAATATGAAGATATTAGACAGGACC	2821
Db	2985	ATCTAAGAGTGAATATCTTTGAGGAATCATGGAATATGAAGATATTAGACAGGACC	3044
Qy	2822	ATTATGAGTCTCTCAAGGAGAGGAGATACCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT	2881
Db	3045	ATTATGAGTCTCTCAAGGAGAGGAGATACCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT	3104
Qy	2882	TCCAAATGGATTTGGCTGTCTGAACCTCACCAGTGAAGCCACGTTTATTGGGACCCAGG	2941
Db	3105	TCCAAATGGATTTGGCTGTCTGAACCTCACCAGTGAAGCCACGTTTATTGGGACCCAGG	3164
Qy	2942	TCCTTGAAGACTATGACTCGCAGAGCTGTGTGACTACATTTGACTGGAAGCTTTCTTTG	3001
Db	3165	TCCTTGAAGACTATGACTCGCAGAGCTGTGTGACTACATTTGACTGGAAGCTTTCTTTG	3224
Qy	3002	ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCCAGGCTTCCCACAGATATTTAAACGACA	3061
Db	3225	ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCCAGGCTTCCCACAGATATTTAAACGACA	3284
Qy	3062	AAACAGTAGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACAC	3121
Db	3285	AAACAGTAGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACAC	3344
Qy	3122	TGATTTAGTCAAAAAGAACTCCGGGGCCGGGGTGTGGTGGTTCCTGCGCAGCAGAGTA	3181
Db	3345	TGATTTAGTCAAAAAGAACTCCGGGGCCGGGGTGTGGTGGTTCCTGCGCAGCAGAGTA	3404
Qy	3182	TCCAAAGACGACATTCACCTGTGTCAGAGGCTGTGTGCCCGCCAGGCTGACAGAGCCATAG	3241
Db	3405	TCCAAAGACGACATTCACCTGTGTCAGAGGCTGTGTGCCCGCCAGGCTGACAGAGCCATAG	3464
Qy	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTGTGCCAGCACGAGGCCATACCT	3301
Db	3465	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTGTGCCAGCACGAGGCCATACCT	3524
Qy	3302	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCCTGCGCATCCGTGACTACTCGGGCCCTGT	3361
Db	3525	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCCTGCGCATCCGTGACTACTCGGGCCCTGT	3584
Qy	3362	TTGCCGTTGCTGCTTTGGGGTAGAAGAGCTGAGAGGCTATGAGGATGATGGTGACG	3421
Db	3585	TTGCCGTTGCTGCTTTGGGGTAGAAGAGCTGAGAGGCTATGAGGATGATGGTGACG	3644
Qy	3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTTCAGAGAG	3481
Db	3645	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTTCAGAGAG	3704
Qy	3482	AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCTGGAGC	3541
Db	3705	AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCTGGAGC	3764
Qy	3542	TCCGAGACCTGCGAAGTTGGGGTACAAAGGCACTCCCGCGGCTCCCTGGCTACCCAGCC	3601
Db	3765	TCCGAGACCTGCGAAGTTGGGGTACAAAGGCACTCCCGCGGCTCCCTGGCTACCCAGCC	3824
Qy	3602	AGCCCGACACACCCGAGAGCTCAATGTGGAGACTCGCAGACATCCAGCAGTCTACAG	3661
Db	3825	AGCCCGACACACCCGAGAGCTCAATGTGGAGACTCGCAGACATCCAGCAGTCTACAG	3884
Qy	3662	GCATTAGGTTAAACAGAAATATTAGCAATGGCACTGCTGTTGAGCAGTCTCAGGCTCTACT	3721
Db	3885	GCATTAGGTTAAACAGAAATATTAGCAATGGCACTGCTGTTGAGCAGTCTCAGGCTCTACT	3944
Qy	3722	TCCTCAATTTGAGTCCAAATATTTGCTGTGGGAGATTTCCAAAGATCAGGTTGAGG	3781
Db	3945	TCCTCAATTTGAGTCCAAATATTTGCTGTGGGAGATTTCCAAAGATCAGGTTGAGG	4004
Qy	3782	ATTATGCAATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTGGACCCATTT	3841
Db	4005	ATTATGCAATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTGGACCCATTT	4064
Qy	3842	TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTATCTTGTGATGATCCT	3901

Db 4065 TGGGATATGATACAGACTAAC- -TTTTTTTTTTTTTGGCTTTTATTTCTTGATGATCCT 4122  
Qy  
Db 3902 CAAGGAATACAACTAG 3919  
Db 4123 CAAGGAATACAACTAG 4140  
RESULT 3  
AEA63658  
ID AEA63658 standard; DNA; 7122 BP.  
XX AC AEA63658;  
XX DT 08-SEP-2005 (first entry)  
XX DE Methionine synthase, MTR, DNA sequence, SEQ ID 2.  
XX KW Neuroleptic; Nootropic; Tranquilizer; folate metabolism;  
XX KW psychiatric disorder; schizophrenia; autism;  
XX KW attention deficit hyperactivity disorder; obsessive-compulsive disorder;  
XX KW gene; ds; Methionine synthase; chromosome 1.  
XX OS Homo sapiens.  
XX PN US6912492-B1.  
XX PD 28-JUN-2005.  
XX PF 23-MAY-2000; 2000US-00577266.  
XX PR 25-MAY-1999; 99US-0136198P.  
XX PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
XX PI Johnson WG, Stenroos ES;  
XX DR WPI; 2005-442724/45.  
PT Estimating the susceptibility of an individual to have offspring that  
PT develop a developmental disorder comprises analyzing the nucleic acids  
PT and/or proteins for genes involved in folate, pyridoxine, and/or  
PT cobalamin metabolism.  
XX Disclosure; SEQ ID NO 2; 84pp; English.  
XX The present invention relates to a method (M1) for estimating the  
XX susceptibility of an individual to have offspring that develop a  
XX developmental disorder. The method comprises analyzing the nucleic acids  
XX and/or proteins from the biological sample, where analyzing results in a  
XX partial or full genotype for the alleles of two or more genes involved in  
XX folate, pyridoxine, and/or cobalamin metabolism. The present sequence is  
XX one such gene involved in folate metabolism. The method is also useful  
XX for diagnosing, preventing, and treating developmental disorders, e.g.  
XX schizophrenia, autism, attention deficit hyperactivity disorder, or  
XX obsessive-compulsive disorder. The protein encoded by the present  
XX sequence is also known as 5-methyltetrahydrofolate-homocysteine S-  
XX methyltransferase. The present sequence is located on chromosome 1q43.  
SQ Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;  
Query Match 99.1%; Score 3884.8; DB 14; Length 7122;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3904; Conservative 0; Mismatches 12; Indels 2; Gaps 1;  
Qy 2 GTCACCTGTGGAGACAGCTCTTCTCGCGGCCCTCTGCGCAAGGAGGACTCGACA 61  
Db 225 GTCACCTGTGGAGACAGCTCTTCTCGCGGCCCTCTGCGCAAGGAGGACTCGACA 284  
Qy 62 ACATGTCACCCCGCTTCCAAGACCTGTGCGAACCCGGAAGGCTGGAAGAAAACCTTCGCGG 121  
Db 285 ACATGTCACCCCGCTTCCAAGACCTGTGCGAACCCGGAAGGCTGGAAGAAAACCTTCGCGG 344

Qy 122 ATGATGATCAATGTCATTTCTGAGAGAGGATTATGTTGCTGGATGGAGGATGGGACCA 181  
Db 345 ATGATGATCAATGTCATTTCTGAGAGAGGATTATGTTGCTGGATGGAGGATGGGACCA 404  
Qy 182 TGATCCAGCGGAGAGCTAAACGAGAACACATTCGAGGTGAGGAATTTAAGATCATG 241  
Db 405 TGATCCAGCGGAGAGCTAAACGAGAACACATTCGAGGTGAGGAATTTAAGATCATG 464  
Qy 242 CCAGCGCGCTGAAAGGCAACATGACATTTTAAAGTATAAATCTCAGCCCTGATGCTATTACC 301  
Db 465 CCAGCGCGCTGAAAGGCAACATGACATTTTAAAGTATAAATCTCAGCCCTGATGCTATTACC 524  
Qy 302 AAATCCATAAGGAATACTTCTGCTGGCTGGGCGAGATATCATTTGAAACAAATACATTTTAGCA 361  
Db 525 AAATCCATAAGGAATACTTCTGCTGGCTGGGCGAGATATCATTTGAAACAAATACATTTTAGCA 584  
Qy 362 GCACTAGTATGCCCCAAGCTGACTATGGCCCTTGAAACACATTTGGGCTACCGGATGAACATGT 421  
Db 585 GCACTAGTATGCCCCAAGCTGACTATGGCCCTTGAAACACATTTGGGCTACCGGATGAACATGT 644  
Qy 422 GCTCTGAGGAGTGGCCAGAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 481  
Db 645 GCTCTGAGGAGTGGCCAGAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 704  
Qy 482 GGTGTTGGCAGGGGCTCTGGGTCGAGCTAAATAAGACACATCTCTGTCTCCCATCTCTGG 541  
Db 705 GGTGTTGGCAGGGGCTCTGGGTCGAGCTAAATAAGACACATCTCTGTCTCCCATCTCTGG 764  
Qy 542 AAAGCCCGGATTAAGGAACATCACTTGTAGAGCTTGTGAAGCATACCAAGAGCAGG 601  
Db 765 AAAGCCCGGATTAAGGAACATCACTTGTAGAGCTTGTGAAGCATACCAAGAGCAGG 824  
Qy 602 CCAAGGACTCTCGATGGCGGGTTCATCTTACTCATTTGAACTATTTTGTGATCTG 661  
Db 825 CCAAGGACTCTCGATGGCGGGTTCATCTTACTCATTTGAACTATTTTGTGATCTG 884  
Qy 662 CCAATGCAAGGCGAGCTTGTGTCACCTCCAAATCTTTTGGAGGAGAAATATCTCCCC 721  
Db 885 CCAATGCAAGGCGAGCTTGTGTCACCTCCAAATCTTTTGGAGGAGAAATATCTCCCC 944  
Qy 722 GGCTATCTTTATTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGACAGA 781  
Db 945 GGCTATCTTTATTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGACAGA 1004  
Qy 782 CAGGAGAGGATTTGTCATCAGCGGTGTCATGAGAACCACTCTGCTGATGGATTAATT 841  
Db 1005 CAGGAGAGGATTTGTCATCAGCGGTGTCATGAGAACCACTCTGCTGATGGATTAATT 1064  
Qy 842 GTGCTTTGGGTGACGCTGAGATGAGACCTTTTATGAAATAATGGAAATGTACAACAG 901  
Db 1065 GTGCTTTGGGTGACGCTGAGATGAGACCTTTTATGAAATAATGGAAATGTACAACAG 1124  
Qy 902 CCTATGCTCTCTGTTATCCCAATCAGGTCTTCCCAACACCTTTGGTGAATGATGAAA 961  
Db 1125 CCTATGCTCTCTGTTATCCCAATCAGGTCTTCCCAACACCTTTGGTGAATGATGAAA 1184  
Qy 962 CGCTTCTATGATGGCCAGACCTTAAAGATTTTGCTATGGATGGCTTGTGCTAATATAG 1021  
Db 1185 CGCTTCTATGATGGCCAGACCTTAAAGATTTTGCTATGGATGGCTTGTGCTAATATAG 1244  
Qy 1022 TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1245 TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1304  
Qy 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTGTC 1141  
Db 1305 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTGTC 1364  
Qy 1142 TAGAGCCCTTCAGGATTTGGACCGTACCAACTTTGTTTAAATTTGGAGAGCGCTGTAATG 1201  
Db 1365 TAGAGCCCTTCAGGATTTGGACCGTACCAACTTTGTTTAAATTTGGAGAGCGCTGTAATG 1424  
Qy 1202 TTGAGGATCAAGGAGTTTGTCTAAACTCATCATGGGAGGAAACTATGAAGAGCCTTGT 1261



|||||  
1425 TTGAGGATCAAGGAATTTGCTAAACTCATCTGCGAGAAACTATGAAGAAGCCCTTGT 1484  
QY  
1262 GTGTTGCAAAAGTGCAGTGGAAATGGAGCCAGGTGTTGGATGTCAAAATGATGATG 1321  
Db  
1485 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCAGGTGTTGGATGTCAAAATGATGATG 1544  
QY  
1322 GCATGCTAGATGCTCAAGTGCATGACCAAGATTTTGCATTAATTTGCTTCCAGCCAG 1381  
Db  
1545 GCATGCTAGATGCTCAAGTGCATGACCAAGATTTTGCATTAATTTGCTTCCAGCCAG 1604  
QY  
1382 ACATCGCAAAAGGTACCTTTGTGATCGACCTCTCCAAATTTTGTGATTTGAAGCTGGGT 1441  
Db  
1605 ACATCGCAAAAGGTACCTTTGTGATCGACCTCTCCAAATTTTGTGATTTGAAGCTGGGT 1664  
QY  
1442 TAAAGTCTGCTCCAGGGAAGTGCATTTGCTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCG 1501  
Db  
1665 TAAAGTCTGCTCCAGGGAAGTGCATTTGCTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCG 1724  
QY  
1502 ACTTCTTTGGAGAGGCCAGGAAGATTTAAAGTATAGGAGTGTCTATGCTGATGCTTATGCTT 1561  
Db  
1725 ACTTCTTTGGAGAGGCCAGGAAGATTTAAAGTATAGGAGTGTCTATGCTGATGCTTATGCTT 1784  
QY  
1562 TTGATGAAGAGGACAGCAGACAGAACAGACACAAATAATCAGAGTGTGACCCCGGCCCT 1621  
Db  
1785 TTGATGAAGAGGACAGCAGACAGAACAGACACAAATAATCAGAGTGTGACCCCGGCCCT 1844  
QY  
1622 ACCATCTGCTTGTGAATAAATCTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1681  
Db  
1845 ACCATCTGCTTGTGAATAAATCTGGGCTTTAATCCNAATGACATTTATTTTGAACCTTAATA 1904  
QY  
1682 TCCTAACCATTTGGGATGGAATGGAGAAACAACTTTGTATGCAATTAATTTTATCCATG 1741  
Db  
1905 TCCTAACCATTTGGGATGGAATGGAGAAACAACTTTGTATGCCATTAATTTTATCCATG 1964  
QY  
1742 CAACAAAGTCATTAAGAAACATTAATCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACT 1801  
Db  
1965 CAACAAAGTCATTAAGAAACATTAATCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACT 2024  
QY  
1802 TGTCTCTTCTCTCCGAGGAATGAAGCCATTTGAGAGCAATGCAATGGGCTTTTCCCTT 1861  
Db  
2025 TGTCTCTTCTCTCCGAGGAATGAAGCCATTTGAGAGCAATGCAATGGGCTTTTCCCTT 2084  
QY  
1862 ACCATGCAATCAAGTCTGGCATGGAATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT 1921  
Db  
2085 ACCATGCAATCAAGTCTGGCATGGAATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT 2144  
QY  
1922 ATGATGATATCCATAAGGAATCTTGCAGCTCTGTGAAGATCTCATCTGGAATTAAGACC 1981  
Db  
2145 ATGATGATATCCATAAGGAATCTTGCAGCTCTGTGAAGATCTCATCTGGAATTAAGACC 2204  
QY  
1982 CTGAGGCCACTGAGAAGCTCTTACGTTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2041  
Db  
2205 CTGAGGCCACTGAGAAGCTCTTACGTTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2264  
QY  
2042 TCATTTCAGACTGATGAGTGGAGAAATGGCCCTGTGCGAAGAACGCTTTGAGTATGCCCTTG 2101  
Db  
2265 TCATTTCAGACTGATGAGTGGAGAAATGGCCCTGTGCGAAGAACGCTTTGAGTATGCCCTTG 2324  
QY  
2102 TGAAGGGCATTTGAAAAACATATTAATTGAGGATCTGAGGAAGCCAGGTTTAAACCAAAAAA 2161  
Db  
2325 TGAAGGGCATTTGAAAAACATATTAATTGAGGATCTGAGGAAGCCAGGTTTAAACCAAAAAA 2384  
QY  
2162 AATATCCCGACCTCTCAATATTAATTGAGGACCCCTGATGAATGGAATGGAATTCGTTG 2221  
Db  
2385 AATATCCCGACCTCTCAATATTAATTGAGGACCCCTGATGAATGGAATGGAATTCGTTG 2444  
QY  
2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTTACCTCAGGTTTATAAGTTCAGCCCGGGTTA 2281  
Db  
2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTTACCTCAGGTTTATAAGTTCAGCCCGGGTTA 2504  
QY  
2282 TGAAGAGAGGCTGTTGGGCACTTTATCCCTTTTATGGAATAAGAAAGAAACCAAGAG 2341  
Db  
|||||

2505 TGAAGAGGCTGTTGGCCACCTTTATCCCTTTTCATGGAATAAGAAAGAGAAACCCAGAG 2564  
QY  
2342 TGCCTTAACCGCACAGTAGAAGAGAGGACCTTTTACAGGGCACCAATCGTGTGCGCCACTG 2401  
Db  
2565 TGCCTTAACCGCACAGTAGAAGAGAGGACCTTTTACAGGGCACCAATCGTGTGCGCCACTG 2624  
QY  
2402 TTAAGAGCGACGTGACGACATAGGCAAGAACATAGTGTGGAGTAGTCTTGGCTGCAATA 2461  
Db  
2625 TTAAGAGCGACGTGACGACATAGGCAAGAACATAGTGTGGAGTAGTCTTGGCTGCAATA 2684  
QY  
2462 ATTTCCGAGTATTTAGATTTTAGGAGTCAATGACTTCATGTGATAGATACCTGAAAGCTGCTC 2521  
Db  
2685 ATTTCCGAGTATTTAGATTTTAGGAGTCAATGACTTCATGTGATAGATACCTGAAAGCTGCTC 2744  
QY  
2522 TTGACCAAAAGCAGATATAAATTTGGCTGTGAGGATCTATCATCTCTTCCCTGGATGAAA 2581  
Db  
2745 TTGACCAAAAGCAGATATAAATTTGGCTGTGAGGATCTATCATCTCTTCCCTGGATGAAA 2804  
QY  
2582 TGAATTTTGTGTCAGAGGAATGAGAGATTTAGCTATAAGGATTTCCATTTGTTGATTTGGAG 2641  
Db  
2805 TGAATTTTGTGTCAGAGGAATGAGAGATTTAGCTATAAGGATTTCCATTTGTTGATTTGGAG 2864  
QY  
2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTGCACCTG 2701  
Db  
2865 GAGCAACCACTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTGCACCTG 2924  
QY  
2702 TAAATCATGCTCTCGGACGCTCCAGAGTGTGGTGTGTTCCAGCTGTTTATAGATGAAA 2761  
Db  
2925 TAAATCATGCTCTCGGACGCTCCAGAGTGTGGTGTGTTCCAGCTGTTTATAGATGAAA 2984  
QY  
2762 ATCTAAAGATGATTAATTTTGAAGGAATCATGGAAGAAATATGAAGATATTTAGACAGACC 2821  
Db  
2985 ATCTAAAGATGATTAATTTTGAAGGAATCATGGAAGAAATATGAAGATATTTAGACAGACC 3044  
QY  
2822 ATTTACAGTCTCTCAAGGAGAGGAGATATTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT 2881  
Db  
3045 ATTTACAGTCTCTCAAGGAGAGGAGATATTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT 3104  
QY  
2882 TCCAAATGGAATGGCTGTCTGAACTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG 2941  
Db  
3105 TCCAAATGGAATGGCTGTCTGAACTCACCCAGTGAAGCCACGTTTATTTGGGACCCAGG 3164  
QY  
2942 TCTTTGAAGACTATGACTGACAGAGCTGGTGAATACATTTGACTGGAAGCCCTTTCTTTG 3001  
Db  
3165 TCTTTGAAGACTATGACTGACAGAGCTGGTGAATACATTTGACTGGAAGCCCTTTCTTTG 3224  
QY  
3002 ATGCTGCGCAGCTCCGCGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTTAAACGACA 3061  
Db  
3225 ATGCTGCGCAGCTCCGCGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTTAAACGACA 3284  
QY  
3062 AAAACAGTAGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACAC 3121  
Db  
3285 AAAACAGTAGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACAC 3344  
QY  
3122 TGAATTAATCAAAAAAATCTCCGGGCCCGGGGTGTTGGTGTCTGGCAGACAGAGTA 3181  
Db  
3345 TGAATTAATCAAAAAAATCTCCGGGCCCGGGGTGTTGGTGTCTGGCAGACAGAGTA 3404  
QY  
3182 TCCAAAGCAGCATTTACCTGTACGAGAGGCTGCTGTGCCCGCCAGGCTGCAAGCCCATAG 3241  
Db  
3405 TCCAAAGCAGCATTTACCTGTACGAGAGGCTGCTGTGCCCGCCAGGCTGCAAGCCCATAG 3464  
QY  
3242 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGAATCTCTGCGACGACGAGGCCATACT 3301  
Db  
3465 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGAATCTCTGCGACGACGAGGCCATACT 3524  
QY  
3302 ACTGCTCTCAGACTTCAATGCTTCCCTTGCATTTCTGGCATTCCTGACTACTCTGGGCTGT 3361  
Db  
3525 ACTGCTCTCAGACTTCAATGCTTCCCTTGCATTTCTGGCATTCCTGACTACTCTGGGCTGT 3584  
QY  
3362 TTGCCGTTGCTCTCTTTGGGTTAGAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3421  
Db  
3585 TTGCCGTTGCTCTCTTTGGGTTAGAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3644



QY 3422 ACTACAGCAGCATCATGTCAAGGCGCTGGGGGACCGGCTGGCAGAGCCCTTTTGAGAG 3481  
DB 3645 ACTACAGCAGCATCATGTCAAGGCGCTGGGGGACCGGCTGGCAGAGCCCTTTTGAGAG 3704  
QY 3482 AGCTCCATGAAGAGTTCCCGCAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541  
DB 3705 AGCTCCATGAAGAGTTCCCGCAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3764  
QY 3542 TCGCAGACCTCGAAGGTTGGGTACAGGGGCATCGGCCGCTCTGGCTACCCAGCC 3601  
DB 3765 TCGCAGACCTCGCAGGCTCGGTACAGGGGCATCGGCCGCTCTGGCTACCCAGCC 3824  
QY 3602 AGCCCGACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3661  
DB 3825 AGCCCGACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3884  
QY 3662 GCATTAGGTTAACAGAACTATTAGCAATGGCACTGTCTTCAGCAGTCTCAGGCTCTACT 3721  
DB 3885 GCATTAGGTTAACAGAACTATTAGCAATGGCACTGTCTTCAGCAGTCTCAGGCTCTACT 3944  
QY 3722 TCTCAATTTGAAGTCCAAATATTTTGTCTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 3781  
DB 3945 TCTCAATTTGAAGTCCAAATATTTTGTCTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 4004  
QY 3782 ATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTCGACCCATT 3841  
DB 4005 ATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTCGACCCATT 4064  
QY 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTATCTTGATGATCCT 3901  
DB 4065 TGGGATATGATACAGACTAAC--TTTTTTTTTTTTTTTGGCTTTTATCTTGATGATCCT 4122  
QY 3902 CAAGGAATACAACTAG 3919  
DB 4123 CAAGGAATACAACTAG 4140

## RESULT 4

ID AAF31101 standard; cDNA; 7224 BP.

XX AC AAF31101;

XX DT 27-APR-2001 (first entry)

XX DE Methionine synthase coding sequence #2.

XX KW Analyte-binding enzyme; analyte analysis; ss.

XX OS Homo sapiens.

XX PN WO200102600-A2.

XX PD 11-JAN-2001.

XX PF 30-JUN-2000; 2000WO-US018057.

XX PR 06-JUL-1999; 99US-00347878.

XX PR 06-DEC-1999; 99US-00457205.

XX PA (GEAT ) GEN ATOMICS.

XX PI Yuan C;

XX DR WPI; 2001-071583/08.

XX PT Assaying method, useful for prognosis and diagnosis of disease, comprises

XX PT contacting sample with a mutant analyte-binding enzyme and detecting

XX PT binding.

XX PS Disclosure; Page; 187pp; English.

XX XX

CC The present invention relates to a method for assaying an analyte in a  
CC sample comprising: contacting the sample with a mutant analyte-binding  
CC enzyme which has binding affinity for the analyte or an immediate analyte  
CC enzymatic conversion product but has attenuated catalytic activity; and  
CC detecting resulting binding. The method is useful in monitoring  
CC biological systems/processes, or prognosis/diagnosis of disease caused by  
CC imbalances of the analytes. The present sequence is a coding sequence  
CC used in the present invention. Note: the present sequence is not shown in  
CC the specification, but was from Genbank, using information given in the  
CC specification

XX SQ Sequence 7224 BP; 2108 A; 1491 C; 1761 G; 1864 T; 0 U; 0 Other;

Query Match 99.1%; Score 3883.2; DB 4; Length 7224;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 3903; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 2 GTCACTGTGGAGAGCAGTCTTCTCTGCCGCGCCCTCTCGCAAGGAGAGATCTCGACA 61  
DB 333 GTCACTGTGGAGAGCAGTCTTCTCTGCCGCGCCCTCTCGCAAGGAGAGATCTCGACA 392  
QY 62 ACATGTCACTGTGGAGAGCAGTCTTCTCTGCCGCGCCCTCTCGCAAGGAGAGATCTCG 121  
DB 393 ACATGTCACTGTGGAGAGCAGTCTTCTCTGCCGCGCCCTCTCGCAAGGAGAGATCTCG 452  
QY 122 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTTATGTCTGGATGGAGGATGGGACCA 181  
DB 453 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTTATGTCTGGATGGAGGATGGGACCA 512  
QY 182 TGATCCAGCGGGAGAGCTAAACGAGAACACACTTCCGAGGTTCAGGAATTTAAAGATCATG 241  
DB 513 TGATCCAGCGGGAGAGCTAAACGAGAACACACTTCCGAGGTTCAGGAATTTAAAGATCATG 572  
QY 242 CCAGCGCGCTGAAAGGCAACAATGACATTTTAAAGTATAAATCAGCCTGATGTCTATTACC 301  
DB 573 CCAGCGCGCTGAAAGGCAACAATGACATTTTAAAGTATAAATCAGCCTGATGTCTATTACC 632  
QY 302 AATCCATTAAGGATATCTTCTGCTGGCGGAGATATCATTTGAAACAAATATCTTTAGCA 361  
DB 633 AATCCATTAAGGATATCTTCTGCTGGCGGAGATATCATTTGAAACAAATATCTTTAGCA 692  
QY 362 GCATAGTATTTGCCCAAGCTGACTATGCGCCCTTGAACACTTTGGCCTACCGGATGAACATGT 421  
DB 693 GCATAGTATTTGCCCAAGCTGACTATGCGCCCTTGAACACTTTGGCCTACCGGATGAACATGT 752  
QY 422 GCTCTGAGAGTGCGCCAGAAAAGCTGCGAGGAGGTAATCTCTCCAGACAGGAATTAAGA 481  
DB 753 GCTCTGAGAGTGCGCCAGAAAAGCTGCGAGGAGGTAATCTCTCCAGACAGGAATTAAGA 812  
QY 482 GGTCTGTGGCAGGGGCTCTGGGCTCCGACTAATAGACACTCTCTGTGTCCCATCTGTGG 541  
DB 813 GGTCTGTGGCAGGGGCTCTGGGCTCCGACTAATAGACACTCTCTGTGTCCCATCTGTGG 872  
QY 542 AAAGCGCGGATTATAGGAACATCACTTTGATGAGCTTTGTAAGCATATCAAGAGCAGG 601  
DB 873 AAAGCGCGGATTATAGGAACATCACTTTGATGAGCTTTGTAAGCATATCAAGAGCAGG 932  
QY 602 CCAAAGGACTTCTCGATGGGGGTTGATATCTTACTCATTTGATAACTATTTTGTGATCTG 661  
DB 933 CCAAAGGACTTCTCGATGGGGGTTGATATCTTACTCATTTGATAACTATTTTGTGATCTG 992  
QY 662 CCAATGCCAGGAGCCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGTCCTCC 721  
DB 993 CCAATGCCAGGAGCCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGTCCTCC 1052  
QY 722 GGCTATCTTTATTTTCAAGGAGCAGTCTGTTGATAAAAGTGGCGGACTCTTTCCGAGACA 781  
DB 1053 GGCTATCTTTATTTTCAAGGAGCAGTCTGTTGATAAAAGTGGCGGACTCTTTCCGAGACA 1112  
QY 782 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGGATTAAT 841  
DB 1113 CAGGAGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGGATTAAT 1172

QY 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTGAAATAATTTGGAATAATGTATCAACAG 901  
DB 1173 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTGAAATAATTTGGAATAATGTATCAACAG 1232  
QY 902 CCTATGCTCTCTGTTATCCCAATCAGAGTCTTCCCAACACACCCCTTTGGTGACCTATGATGAAA 961  
DB 1233 CCTATGCTCTCTGTTATCCCAATCAGAGTCTTCCCAACACACCCCTTTGGTGACCTATGATGAAA 1292  
QY 962 GGCCTTCTATGATGGCCAGACCTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
DB 1293 GGCCTTCTATGATGGCCAGACCTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1352  
QY 1022 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAAGCTGTGAAA 1081  
DB 1353 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAAGCTGTGAAA 1412  
QY 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGCTGCTGTC 1141  
DB 1413 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGCTGCTGTC 1472  
QY 1142 TAGAGCCCTTTCAGGATTTGGACCTGACCACTTTTGTATGATGGAGCGCTGTAAATG 1201  
DB 1473 TAGAGCCCTTTCAGGATTTGGACCTGACCACTTTTGTATGATGGAGCGCTGTAAATG 1532  
QY 1202 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAAGCTGTGAAA 1261  
DB 1533 TTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAAGCTGTGAAA 1592  
QY 1262 GTGTGCAAAAGTGCAGTGGAAATGGAGCCAGGTTTGGATGTCAAATGATGATGATG 1321  
DB 1593 GTGTGCAAAAGTGCAGTGGAAATGGAGCCAGGTTTGGATGTCAAATGATGATGATG 1652  
QY 1322 GCATGCTAGATGCTGCAAGTGCATGCAATGATGCAATGATGCAATGATGCAATGATG 1381  
DB 1653 GCATGCTAGATGCTGCAAGTGCATGCAATGATGCAATGATGCAATGATGCAATGATG 1712  
QY 1382 ACATGCAAAAGTGCATGCTGCAATGATGCAATGATGCAATGATGCAATGATGCAATG 1441  
DB 1713 ACATGCAAAAGTGCATGCTGCAATGATGCAATGATGCAATGATGCAATGATGCAATG 1772  
QY 1442 TAAAGTCTGCTGCAAGGAGTGCATGCTGCAATGATGCAATGATGCAATGATGCAATG 1501  
DB 1773 TAAAGTCTGCTGCAAGGAGTGCATGCTGCAATGATGCAATGATGCAATGATGCAATG 1832  
QY 1502 ACTTCTTTGGAGAGCCAGGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1561  
DB 1833 ACTTCTTTGGAGAGCCAGGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1892  
QY 1562 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCAACCCGGGCT 1621  
DB 1893 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCAACCCGGGCT 1952  
QY 1622 ACCATCTGCTTGTGAAAATACTGGGCTTTTAAATCAATGATGATTTTGTGACCTTAATA 1681  
DB 1953 ACCATCTGCTTGTGAAAATACTGGGCTTTTAAATCAATGATGATTTTGTGACCTTAATA 2012  
QY 1682 TCCTAACCATTTGGGACTTGGATGAGGAGAACACACTTTGTATGCCATTAATTTTATCCATG 1741  
DB 2013 TCCTAACCATTTGGGACTTGGATGAGGAGAACACACTTTGTATGCCATTAATTTTATCCATG 2072  
QY 1742 CAACAAAGTCAATTAAGAAACATTAACCTGAGGACAGAAATAGTGGAGGCTTTTCCAACT 1801  
DB 2073 CAACAAAGTCAATTAAGAAACATTAACCTGAGGACAGAAATAGTGGAGGCTTTTCCAACT 2132  
QY 1802 TGTCTCTTCTCTCCGAGGATGGAAGCCATTCAGAGCAATGCAATGGGCTTTTCCCTTT 1861  
DB 2133 TGTCTCTTCTCTCCGAGGATGGAAGCCATTCAGAGCAATGCAATGGGCTTTTCCCTTT 2192  
QY 1862 ACCATGCAATCAAGTCTGGCATGACATGGAGATGATGATGCTGGAAACCTCCCTGCTGT 1921  
DB 2193 ACCATGCAATCAAGTCTGGCATGACATGGAGATGATGATGCTGGAAACCTCCCTGCTGT 2252  
QY 1922 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC 1981

DB 2253 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC 2312  
QY 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2041  
DB 2313 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2372  
QY 2042 TCATTGAGCTGATGATGGAGAAATGGCCCTCTGAGAACCCCTTGTAGTATGCCCTTG 2101  
DB 2373 TCATTGAGCTGATGATGGAGAAATGGCCCTCTGAGAACCCCTTGTAGTATGCCCTTG 2432  
QY 2102 TGAAGGGCATTTGAAAACATATTTATGAGGATCTGAGGAAAGCCAGGTAAACCAAAAAA 2161  
DB 2433 TGAAGGGCATTTGAAAACATATTTATGAGGATCTGAGGAAAGCCAGGTAAACCAAAAAA 2492  
QY 2162 AATATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGTGATGAATGGAATTTGTTG 2221  
DB 2493 AATATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGTGATGAATGGAATTTGTTG 2552  
QY 2222 GTGATCTTTTGGAGCTGGAAAATTTTCTACTCAGGTTTAAAGTCAAGCCGGGTTA 2281  
DB 2553 GTGATCTTTTGGAGCTGGAAAATTTTCTACTCAGGTTTAAAGTCAAGCCGGGTTA 2612  
QY 2282 TGAAGAGGCTGTGGCCACCTTTATCCCTTTTCAATGGAAGAAAGAGAAAGAACCCAGAG 2341  
DB 2613 TGAAGAGGCTGTGGCCACCTTTATCCCTTTTCAATGGAAGAAAGAGAAAGAACCCAGAG 2672  
QY 2342 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACCAGGGCACCATCTGCTGCCACTG 2401  
DB 2673 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACCAGGGCACCATCTGCTGCCACTG 2732  
QY 2402 TTAAGGGGAGCTGTGGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2461  
DB 2733 TTAAGGGGAGCTGTGGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2792  
QY 2462 ATTTCCGAGTTATTTAGGAGTCAATGATCCATGATAGATAGATAGTCAAGAGCTGCTC 2521  
DB 2793 ATTTCCGAGTTATTTAGGAGTCAATGATCCATGATAGATAGATAGTCAAGAGCTGCTC 2852  
QY 2522 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2581  
DB 2853 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2912  
QY 2582 TGATTTTGTGTCACAGGAAATGAGAGATTAGCTATAAGGNTTCCATTTGATTTGGAG 2641  
DB 2913 TGATTTTGTGTCACAGGAAATGAGAGATTAGCTATAAGGNTTCCATTTGATTTGGAG 2972  
QY 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAAATAGCTCCGAGATACAGTGACCTG 2701  
DB 2973 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAAATAGCTCCGAGATACAGTGACCTG 3032  
QY 2702 TAAATCCATGCTGTCGAGCGCTCCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2761  
DB 3033 TAAATCCATGCTGTCGAGCGCTCCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3092  
QY 2762 ATCTAAAGGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC 2821  
DB 3093 ATCTAAAGGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC 3152  
QY 2822 ATTTAGAGTCTCTCAAGGAGAGAGATATTACCCCTTTAGTCAAGCCAGAAAAGTGGTT 2881  
DB 3153 ATTTAGAGTCTCTCAAGGAGAGAGATATTACCCCTTTAGTCAAGCCAGAAAAGTGGTT 3212  
QY 2882 TCCAAATGGAATGGCTGTCTGAAACCTTACCAGTGAAGCCACCTTTTATTTGGAGCCAGG 2941  
DB 3213 TCCAAATGGAATGGCTGTCTGAAACCTTACCAGTGAAGCCACCTTTTATTTGGAGCCAGG 3272  
QY 2942 TCTTTGAAGACTATGACCTGCAGAGCTGTGTGACTACATTTGACTGGAAGCCCTTTCTTTG 3001  
DB 3273 TCTTTGAAGACTATGACCTGCAGAGCTGTGTGACTACATTTGACTGGAAGCCCTTTCTTTG 3332  
QY 3002 ATGCTGCGAGCTCCGGGGCAAGTACCCGAAATGAGGCTTCCCAAGATATTTTAAAGACA 3061

3333	ATGCTCGCAGCTCCGGGGCAAGTACCGGAATCGAGGCTTTCCCAAGATATTTAAACGACA	3392	KW	Methionine synthase; human; hyperhomocysteinaemia; neural tube defect; cardiovascular disease; colon cancer; diagnosis; therapy; ss.
3062	AAACAGTAGGTGGAGAGCCAGGAGGCTACGATGATGCCACAAATATGCTGAACACAC	3121	KW	
3393	AAACAGTAGGTGGAGAGCCAGGAGGCTACGATGATGCCACAAATATGCTGAACACAC	3452	OS	Homo sapiens.
3122	TGATTAGTCAAAAGAAATCCGGGCCCGGGTGTGGTTCTGGCCACACAGAGTA	3181	PH	Key
3453	TGATTAGTCAAAAGAAATCCGGGCCCGGGTGTGGTTCTGGCCACACAGAGTA	3512	FT	Location/Qualifiers
3182	TCCAAGACGACATTCCTGTACGAGAGGCTGTGTGCCCGCCAGGCTGCAGAGCCCATAG	3241	FT	64..3861
3513	TCCAAGACGACATTCCTGTACGAGAGGCTGTGTGCCCGCCAGGCTGCAGAGCCCATAG	3572	FT	/*tag= a
3242	CCACTTTCTATGGGTAAAGCAACAGGCTGAGAGGACTCTGCCAGACGAGCCATACT	3301	FT	/EC_number= "2.1.1.13"
3573	CCACTTTCTATGGGTAAAGCAACAGGCTGAGAGGACTCTGCCAGACGAGCCATACT	3632	XX	CA2217153-A.
3302	ACTGCCCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGCCCTGT	3361	XX	27-MAY-1998.
3633	ACTGCCCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGCCCTGT	3692	XX	27-NOV-1997; 97CA-02217153.
3362	TGCGCGTTCCTGCTTTGGGTAGAGAGCTGAGAGGCTTATGAGATGATGGTAGC	3421	XX	27-NOV-1996; 96US-0031964P.
3693	TGCGCGTTCCTGCTTTGGGTAGAGAGCTGAGAGGCTTATGAGATGATGGTAGC	3752	PR	20-JUN-1997; 97US-0050310P.
3422	ACTACAGAGCATCATGTGCAAGCGCTGGGGACCGGCTGGCAGAGGCCCTTTCGAGAAG	3481	XX	(MART-) MARTINEX R & D INC.
3753	ACTACAGAGCATCATGTGCAAGCGCTGGGGACCGGCTGGCAGAGGCCCTTTCGAGAAG	3812	XX	Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
3482	AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG	3541	PI	WPI; 1998-569373/49.
3813	AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG	3872	XX	P-PSDB; AAW68594.
3542	TGCGAGACTCGGAGGCTGGGTACAAAGGCACTCCGCCGGCTCCTGGCTACCCAGCC	3601	PT	DNA encoding methionine synthase polypeptide - and corresponding polypeptide, cells, antibody and therapeutic methods.
3873	TGCGAGACTCGGAGGCTGGGTACAAAGGCACTCCGCCGGCTCCTGGCTACCCAGCC	3932	XX	Claim 6; Fig 3; 64pp; English.
3602	AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3661	XX	This cDNA clone encodes human wild-type methionine synthase (see AAW68594). A survey of the NCBI databases yielded several sequences corresponding to methionine synthases from different organisms.
3933	AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTTGCGAGACATCGAGCAGTCTACAG	3992	CC	Comparison of these sequences generated 4 very conserved regions. Degenerate oligonucleotides (see AAV34064-104) based on these conserved sequences were used as primers for RT-PCR with human and mouse mRNA. PCR products were subcloned, sequenced and aligned. Additional sequences were subsequently obtained by further PCR and inverse PCR. The full-length sequence encodes a 1265-amino acid protein of predicted mol.wt. 141 kDa. Using FISH, the gene was mapped to human chromosome band 1q43. Mutations in this gene are associated with hyperhomocysteinaemia. One missense mutation (2758 C to G) and a 3 bp deletion (2640-42) have been identified in patients of the cblG complementation group. A polymorphism (2756 A to G) has also been detected. The invention relates to the diagnosis and treatment of patients at risk for methionine synthase deficiency, and associated altered risk for neural tube defects, cardiovascular disease, and colon cancer. Methods for detecting sequence variants for methionine synthase, involving SSCP analysis or RFLP polymorphism analysis, are claimed, as are methods for screening for therapeutic compounds that modulate methionine synthase activity
3662	GCATTAGGTTAAAGATCATTAGCAATGGCACCTGCTTCAGCAGTCTCAGGCTCTACT	3721	CC	Sequence 3920 BP; 1114 A; 866 C; 975 G; 965 T; 0 U; 0 Other;
3993	GCATTAGGTTAAAGATCATTAGCAATGGCACCTGCTTCAGCAGTCTCAGGCTCTACT	4052	CC	Query Match
3722	TCTCAATTTGAACTCCAAATATTTTGTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGG	3781	CC	Best Local Similarity 96.8%; Score 3793.8; DB 2; Length 3920;
4053	TCTCAATTTGAACTCCAAATATTTTGTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGG	4112	CC	Best Local Similarity 98.5%; Pred. No. 0;
3782	ATTATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATT	3841	CC	Matches 3863; Conservative 0; Mismatches 57; Indels 3; Gaps 3;
4113	ATTATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATT	4172	CC	
3842	TGGGATATGATACAGACTAACTTTTGTGCTTTTGTGCTTTTATCTTGTGATGATCCT	3901	CC	
4173	TGGGATATGATACAGACTAAC--TTTGTGCTTTTGTGCTTTTATCTTGTGATGATCCT	4230	CC	
3902	CAAGGAATACAACTAG 3919		Qy	1
4231	CAAGGAATACAACTAG 4248		Db	1
RESULT 5			Qy	61
AAV34063			Db	61
XX			Qy	121
AAV34063;			Db	121
01-MAR-1999 (first entry)			Qy	181
Human methionine synthase cDNA (wild-type).			Db	181

Qy 241 GCCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTTAC 300  
Db 241 GCCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTTAC 300  
Qy 301 CAATCCATTAAGGAATATCTGCTGGCTGGGCGAGATATCATTTGAACAAAATACATTTTAGC 360  
Db 301 CAATCCATTAAGGAATATCTGCTGGCTGGGCGAGATATCATTTGAACAAAATACATTTTAGC 360  
Qy 361 AGCACTAGTATGCCCCAAGCTGACATATGSCCTTGAACACTTGGCTACCGGATGAACATG 420  
Db 361 AGCACTAGTATGCCCCAAGCTGACATATGSCCTTGAACACTTGGCTACCGGATGAACATG 420  
Qy 421 TGCTCTCAGAGATGGCCAGAAAAGCTGCCAGAGAGGTAACTCTCTCCAGACAGGAATTAAG 480  
Db 421 TGCTCTCAGAGATGGCCAGAAAAGCTGCCAGAGAGGTAACTCTCTCCAGACAGGAATTAAG 480  
Qy 481 AGGTTTGGCAGGGGCTCTGGCTCCGACTTAATAGACACTCTCTGTGTCGCCCATCTGTG 540  
Db 481 AGGTTTGGCAGGGGCTCTGGCTCCGACTTAATAGACACTCTCTGTGTCGCCCATCTGTG 540  
Qy 541 GAAAGCCGGATTAAGGAACATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAG 600  
Db 541 GAAAGCCGGATTAAGGAACATCAATTTGATGAGCTTTGTAAGCATACCAAGAGCAG 600  
Qy 601 GCCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAACCTATTTTGTACT 660  
Db 601 GCCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAACCTATTTTGTACT 660  
Qy 661 GCCAATGCCAAGGCGCTCTGTCCTCCACTCAAAATCTTTTGGAGGAATATGCTCCC 720  
Db 661 GCCAATGCCAAGGCGCTCTGTCCTCCACTCAAAATCTTTTGGAGGAATATGCTCCC 720  
Qy 721 CGGCTATCTTTATTTCAAGGACCATCTGTATGATAAAGTGGCGGACTCTTTCCGAGCAG 780  
Db 721 CGGCTATCTTTATTTCAAGGACCATCTGTATGATAAAGTGGCGGACTCTTTCCGAGCAG 780  
Qy 781 ACAGAGAGGATTTGTCAATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGAATTAAT 840  
Db 781 ACAGAGAGGATTTGTCAATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGAATTAAT 840  
Qy 841 TGTCCTTTGGGTGAGCTGAGATGAGACTTTTATTTGAATTAATTTGGAATGTGACAA 900  
Db 841 TGTCCTTTGGGTGAGCTGAGATGAGACTTTTATTTGAATTAATTTGGAATGTGACAA 900  
Qy 901 GCCTATGTCTCTGTTATCCCAATGCAGGTCTTCCCAACACTTTTGGTGACTATGATGAA 960  
Db 901 GCCTATGTCTCTGTTATCCCAATGCAGGTCTTCCCAACACTTTTGGTGACTATGATGAA 960  
Qy 961 ACGCCCTTCTATGATGGCCCAAGCACCTAAAGGATTTTGTATGATGGCTTGGTCAATATA 1020  
Db 961 ACCCTTCTATGATGGCCCAAGCACCTAAAGGATTTTGTATGATGGCTTGGTCAATATA 1020  
Qy 1021 GTTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1080  
Db 1021 GTTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1080  
Qy 1081 AATTGTAAGCTTAGAGTTCCACCTGCACTCTTTTGAAGGACATATGTTACTGTCTGTT 1140  
Db 1081 AATTGTAAGCTTAGAGTTCCACCTGCACTCTTTTGAAGGACATATGTTACTGTCTGTT 1140  
Qy 1141 CTAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGTAAATTTGGAGAGCGCTGTAAT 1200  
Db 1141 CTAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGTAAATTTGGAGAGCGCTGTAAT 1200  
Qy 1201 GTTGAGAGATCAAGGAATTTGCTTAACTCATCATGCGGAGAACTATGAAGAGCCCTTG 1260  
Db 1201 GTTGAGAGATCAAGGAATTTGCTTAACTCATCATGCGGAGAACTATGAAGAGCCCTTG 1260  
Qy 1261 TGCTGTCCAAAGTGCAGGTGGAATGGAGCCCAAGGTGTTGATGTGTCACATGATGAT 1320  
Db 1261 TGCTGTCCAAAGTGCAGGTGGAATGGAGCCCAAGGTGTTGATGTGTCACATGATGAT 1320

Qy 1321 GGCAATGCTAGATGCTCAAGTGCACATGACAGATTTTGCACCTTAATTTGCTTCCGAGCA 1380  
Db 1321 GGCAATGCTAGATGCTCAAGTGCACATGACAGATTTTGCACCTTAATTTGCTTCCGAGCA 1380  
Qy 1381 GACATCCCAAGGTGACTTTTGTGCTCATGACTCTCTCCAAATTTGCTGATTTGAAGCTGGG 1440  
Db 1381 GACATCCCAAGGTGACTTTTGTGCTCATGACTCTCTCCAAATTTGCTGATTTGAAGCTGGG 1440  
Qy 1441 TTTAAAGTCTGCCAAGGAAAGTGTCAATAGCATTTAGTCTGAAGGAGGAGAGGAC 1500  
Db 1441 TTTAAAGTCTGCCAAGGAAAGTGTCAATAGCATTTAGTCTGAAGGAGGAGAGGAC 1500  
Qy 1501 GACTTCTTGAGAGGCGCCAGGAAGTAAAGATGTGAGCTGCTATGTTGGTCTATGGCT 1560  
Db 1501 GACTTCTTGAGAGGCGCCAGGAAGTAAAGATGTGAGCTGCTATGTTGGTCTATGGCT 1560  
Qy 1561 TTTGATCAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCC 1620  
Db 1561 TTTGATCAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCC 1620  
Qy 1621 TACCATCTCTTGTGAAAAAATCGGCTTTAAATCCAAATGACATTAATTTTCACTTAAT 1680  
Db 1621 TACCATCTCTTGTGAAAAAATCGGCTTTAAATCCAAATGACATTAATTTTCACTTAAT 1680  
Qy 1681 ATCTTAACCATTTGGGACTGGAATGGAGGAAACAACTTGTATGTCATTAATTTTATCCAT 1740  
Db 1681 ATCTTAACCATTTGGGACTGGAATGGAGGAAACAACTTGTATGTCATTAATTTTATCCAT 1740  
Qy 1741 GCAACAAAAGTCAATTAAGAAAAATACCTTGGAGCCAGAAATAGTGAGGTCTTTTCAAC 1800  
Db 1741 GCAACAAAAGTCAATTAAGAAAAATACCTTGGAGCCAGAAATAGTGAGGTCTTTTCAAC 1800  
Qy 1801 TTGTCTCTCTTCCGAGGATGGAGCCATTCGAGAAGCAATGCAATGCGGTTTTCTTT 1860  
Db 1801 TTGTCTCTCTTCCGAGGATGGAGCCATTCGAGAAGCAATGCAATGCGGTTTTCTTT 1860  
Qy 1861 TACCATCAATCAAGTCTGCGCATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTG 1920  
Db 1861 TACCATCAATCAAGTCTGCGCATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTG 1920  
Qy 1921 TATGATCATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAC 1980  
Db 1921 TATGATCATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAC 1980  
Qy 1981 CCTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAA 2040  
Db 1981 CCTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAA 2040  
Qy 2041 GTCAATCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGTAGTATGCCCTT 2100  
Db 2041 GTCAATCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGTAGTATGCCCTT 2100  
Qy 2101 GTGAAGGCAATGGAACATATTTAGGATCTGAGGAGCCAGGTAAACCAAAAA 2160  
Db 2101 GTGAAGGCAATGGAACATATTTAGGATCTGAGGAGCCAGGTAAACCAAAAA 2160  
Qy 2161 AATATATCCCGACTCTCAATATAATTTGAAGGACCCCTGATGAATGAATGAATTTGTT 2220  
Db 2161 AATATATCCCGACTCTCAATATAATTTGAAGGACCCCTGATGAATGAATGAATTTGTT 2220  
Qy 2221 GGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTT 2280  
Db 2221 GGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTT 2280  
Qy 2281 ATGAAGAGGCTGTTGGCCACTTATCCCTTCTGAAAAAAGAGAGAGAACCCAGA 2340  
Db 2281 ATGAAGAGGCTGTTGGCCACTTATCCCTTCTGAAAAAAGAGAGAGAACCCAGA 2340  
Qy 2341 GTGCTTAAAGGCAAGTAGAAGAGAGGACCCCTTACCGAGGACCATCTGCTGCGCACT 2400  
Db 2341 GTGCTTAAAGGCAAGTAGAAGAGAGGACCCCTTACCGAGGACCATCTGCTGCGCACT 2400  
Qy 2401 GTTAAAGGCGAGCTGTCACGATAGGCAAGAACATAGTTGGAGTGTCTTTGGCTGCAAT 2460

Db GTTAAAGCGAGTGCGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAAT 2460  
Qy AATTTCGAGTTATGATTTAGGAGTCACTGATCCATGCTGATAGATGACTGAAAGCTGCT 2520  
Db AATTTCGAGTTATGATTTAGGAGTCACTGATCCATGCTGATAGATGACTGAAAGCTGCT 2520  
Qy CTTCGACCAAGCAGATATAATTTGGCTGTCAGGACTCATCACTCCCTTCCTCGATGAA 2580  
Db CTTCGACCAAGCAGATATAATTTGGCTGTCAGGACTCATCACTCCCTTCCTCGATGAA 2580  
Qy ATGATTTTGTTCGCAAGAAATGAGAGATTAGCTATAAGGATTTCCATTGTTGATGGA 2640  
Db ATGATTTTGTTCGCAAGAAATGAGAGATTAGCTATAAGGATTTCCATTGTTGATGGA 2640  
Qy GGAGCAACCTTCGCAAGAAATGAGAGATTAGCTATAAGGATTTCCATTGTTGATGGA 2700  
Db GGAGCAACCTTCGCAAGAAATGAGAGATTAGCTATAAGGATTTCCATTGTTGATGGA 2700  
Qy GTATCCATGTCCTGGAGCGCTCCAGAGTGTGGTGTCTTCCAGCTGTTAGATGAA 2760  
Db GTATCCATGTCCTGGAGCGCTCCAGAGTGTGGTGTCTTCCAGCTGTTAGATGAA 2760  
Qy AATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGAC 2820  
Db AATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGAC 2820  
Qy CATTATGAGTCTCTCAAGAGAGAGATATCTTACCTTTAAGTCAAGCCAGAAAAGTGGT 2880  
Db CATTATGAGTCTCTCAAGAGAGAGATATCTTACCTTTAAGTCAAGCCAGAAAAGTGGT 2880  
Qy TTCCAAATGGATTTGGCTGCTGTAACCTCAACGAGTGTGGTGTCTTCCAGCTGTTAGATGAA 2940  
Db TTCCAAATGGATTTGGCTGCTGTAACCTCAACGAGTGTGGTGTCTTCCAGCTGTTAGATGAA 2940  
Qy GTCTTTGAGGACTATGACTGTCAGAAAGTGTGGTGTCTTCCAGCTGTTAGATGAA 3000  
Db GTCTTTGAGGACTATGACTGTCAGAAAGTGTGGTGTCTTCCAGCTGTTAGATGAA 3000  
Qy GATGTCCTGGAGTCTCCGGGGAAGTACCCGAAATGAGAGTCTTCCGCAAGATATTAAAGAC 3060  
Db GATGTCCTGGAGTCTCCGGGGAAGTACCCGAAATGAGAGTCTTCCGCAAGATATTAAAGAC 3060  
Qy AAACAGTAGTGGAGAGCCAGGAGTCTACGATGATGCCCAATATGCTGAACACA 3120  
Db AAACAGTAGTGGAGAGCCAGGAGTCTACGATGATGCCCAATATGCTGAACACA 3120  
Qy CTGATTAGTCAAAAGAAACTCCGGGCCGGGTGTGGTGTCTGCGCCAGCACAGAGT 3180  
Db CTGATTAGTCAAAAGAAACTCCGGGCCGGGTGTGGTGTCTGCGCCAGCACAGAGT 3180  
Qy ATCCAAAGACGACATTCACCTGTACGCGAGGCTGTGTGCCAGGCTGCGAGAGCCCAT 3240  
Db ATCCAAAGACGACATTCACCTGTACGCGAGGCTGTGTGCCAGGCTGCGAGAGCCCAT 3240  
Qy GCCACTTCTATGCGTTAAGGCAACAGGCTGAGAGGACTGCGCAGCACGAGCCATAC 3300  
Db GCCACTTCTATGCGTTAAGGCAACAGGCTGAGAGGACTGCGCAGCACGAGCCATAC 3300  
Qy TACTGCTCTCAGACTTCACTGCTCCCTTTCGCAATCTGCGCATCCGTCGCTTCCGCTG 3360  
Db TACTGCTCTCAGACTTCACTGCTCCCTTTCGCAATCTGCGCATCCGTCGCTTCCGCTG 3360  
Qy TTTCGCTGCTGCTTTGGGTTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTAC 3420  
Db TTTCGCTGCTGCTTTGGGTTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTAC 3420  
Qy GACTACAGGATCATGCTCAAGGCTGGGGACCGCTGGCAGAGGCTTTTCAGAA 3480  
Db GACTACAGGATCATGCTCAAGGCTGGGGACCGCTGGCAGAGGCTTTTCAGAA 3480  
Qy GAGCTCCATGAAAGAGTTCCGCCAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGAC 3540

Db 3481 GAGCTCCATGAAGAGTTCCGCAAGAACTGTGGGCTTACTGTGGCAGTGACCAGCTGGAC 3540  
Qy GTTCGAGACCTGCGAGGTTGCGGTACAGGGCATCCGCCGCTCTCTGCTACCCAGC 3600  
Db GTTCGAGACCTGCGAGGTTGCGGTACAGGGCATCCGCCGCTCTCTGCTACCCAGC 3600  
Qy CAGCCGACACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACA 3660  
Db CAGCCGACACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACA 3660  
Qy GGCAATTAGGTTAAACAGAAATCATTTAGCAATGGCACTGCTTTCAGCAGTCTCA-GGCCCTCTA 3719  
Db GGCAATTAGGTTAAACAGAAATCATTTAGCAATGGCACTGCTTTCAGCAGTCTCA-GGCCCTCTA 3720  
Qy CTCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAAGATTTCCAGGATCAGGTTGA 3779  
Db CTCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAAGATTTCCAGGATCAGGTTGA 3779  
Qy GGATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCAT 3839  
Db GGATTATGCAATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCAT 3839  
Qy TTTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGGCCTTTTTTATCTTGATGAT 3898  
Db TTTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGGCCTTTTTTATCTTGATGAT 3899  
Qy CCTCAAGGAAATACACCTAG 3919  
Db CCTCAAGGAAATACACCTAG 3920

RESULT 6  
ADI39109  
ID ADI39109 standard; DNA; 3798 BP.  
XX  
AC ADI39109;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Homo sapiens meth DNA.  
XX  
KW Meth; animal feed additive; sulphur; human nutrition; animal nutrition;  
KW cosmetic; pharmaceutical; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3798  
FT /\*tag= a  
FT /product= "meth"  
XX  
WO2003087386-A2.  
PD 23-OCT-2003.  
XX  
PF 16-APR-2003; 2003WO-EP004010.  
XX  
PR 17-APR-2002; 2002DE-01017058.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Kroeger B, Zelder O, Klopprogge C, Schroeder H, Haefner S;  
XX  
DR WPI; 2003-877106/81.  
XX  
PT Fermentative production of sulfur-containing compounds, particularly L-methionine, useful as feed additives, by using Coryneform bacteria that overexpress methionine synthase.  
XX  
PS Claim 5; SEQ ID NO 43; 304pp; German.  
XX



CC This invention describes a novel method for the fermentative production  
 CC of methionine by growing a sulphur-producing Coryneform bacteria that  
 CC expresses at least one heterologous nucleic acid encoding a protein with  
 CC methionine synthase (Meth) activity. Methionine accumulates in the medium  
 CC or the cells. The method can be used to produce an L-methionine-  
 CC containing animal feed additive by culturing an L-Met-producing  
 CC microorganism, removing water from the resulting broth, removing 0-100%  
 CC of the biomass formed and drying the product to produce the feed additive  
 CC in powdered or granular form. The nucleic acid encoding Meth has sequence  
 CC homology less than 100% with respect to the meth coding sequence of  
 CC Corynebacterium glutamicum ATCC 13032. Optionally at least one other gene  
 CC in the methionine biosynthesis pathway (e.g. aspartate kinase,  
 CC glyceraldehyde-3-phosphate dehydrogenase or 3-phosphoglycerate kinase) is  
 CC also amplified or mutated so that it is not affected by metabolites. Also  
 CC at least one metabolic pathway that reduces production of methionine is  
 CC dehydratase or threonine synthase (e.g. homoserine kinase, threonine  
 CC produce L-methionine, useful as an additive for animal feeds. More  
 CC generally sulphur-containing fine chemicals are useful in human and  
 CC animal nutrition, cosmetics and pharmaceuticals. This sequence encodes a  
 CC meth protein described in the disclosure of the invention.

XX SQ Sequence 3798 BP; 1082 A; 809 C; 982 G; 925 T; 0 U; 0 Other;

Query Match 96.6%; Score 3785.2; DB 10; Length 3798;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3790; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	64	ATGTCACCGCGCTCCAGACCTCTCGCAACCGAGGCTCTGAAGAAACCCCTGCGGGAT	123
Db	1	ATGTCACCGCGCTCCAGACCTCTCGCAACCGAGGCTCTGAAGAAACCCCTGCGGGAT	60
Qy	124	GAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTGTGGATGGAGGATGGGACCATG	183
Db	61	GAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTGTGGATGGAGGATGGGACCATG	120
Qy	184	ATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	243
Db	121	ATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	180
Qy	244	AGGCGCGTGAAGGCAACATGACATTTAAGTATTAACCTCAGCTGATGTCATTTACCA	303
Db	181	AGGCGCGTGAAGGCAACATGACATTTAAGTATTAACCTCAGCTGATGTCATTTACCA	240
Qy	304	ATCCATAAGGAATATCTTGGCTGGCGGAGATATCAATGAACAAATATCTTTTAGCAGC	363
Db	241	ATCCATAAGGAATATCTTGGCTGGCGGAGATATCAATGAACAAATATCTTTTAGCAGC	300
Qy	364	ACTAGTATGCCCAAGCTGACTATGGCCTTGAACACTTGGCCTACCGGATGAACATGTC	423
Db	301	ACTAGTATGCCCAAGCTGACTATGGCCTTGAACACTTGGCCTACCGGATGAACATGTC	360
Qy	424	TCTGCAGAGTGGCCAGAAAGCTGCCGAGAGGATATCTCTCCAGACAGGAATTAAGAGG	483
Db	361	TCTGCAGAGTGGCCAGAAAGCTGCCGAGAGGATATCTCTCCAGACAGGAATTAAGAGG	420
Qy	484	TTTGTGTCAGGGGCTCTGGGTCGAGCTAAAGACACTCTCTGTGTCCTCATCTGTGGAA	543
Db	421	TTTGTGTCAGGGGCTCTGGGTCGAGCTAAAGACACTCTCTGTGTCCTCATCTGTGGAA	480
Qy	544	AGGCGGATTTATAGGAACATCACATTTGATGAGCTTTGTAAGACATACCAAGACGAGCC	603
Db	481	AGGCGGATTTATAGGAACATCACATTTGATGAGCTTTGTAAGACATACCAAGACGAGCC	540
Qy	604	AAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGATCTGCC	663
Db	541	AAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGATCTGCC	600
Qy	664	AATGCCAAGCAGCGCTTTGTCACCTCCAAATCTTTTTCAGGAGGAATATGCTCCCGG	723
Db	601	AATGCCAAGCAGCGCTTTGTCACCTCCAAATCTTTTTCAGGAGGAATATGCTCCCGG	660
Qy	724	CCTATCTTTATTTTCAGGAGCGATCGTTGATATAAAGTGGGGGAGCTCTTTCCGGACAGCA	783

Db	661	CCATATCTTTATTTTCAGGAGCGATCGTTTGATATAAAGTGGCGGAGCTCTTTCCGACAGCA	720
Qy	784	GGAGAGGAGTTTGTCTATCAGCTGTCTCTCATGAGAGAACCACTCTGCAATTTGATTAATTTGT	843
Db	721	GGAGAGGAGTTTGTCTATCAGCTGTCTCTCATGAGAGAACCACTCTGCAATTTGATTAATTTGT	780
Qy	844	GCCTTTGGTGACGCTGATGATGAGACCTTTTATTTGAAATAATTTGGAATAATGTACACAGCC	903
Db	781	GCCTTTGGTGACGCTGATGATGAGACCTTTTATTTGAAATAATTTGGAATAATGTACACAGCC	840
Qy	904	TATGTCTCTGTTTATCCCAATGAGGCTCTTCCCAACACCTTTTGGTGAATGATGAAGACG	963
Db	841	TATGTCTCTGTTTATCCCAATGAGGCTCTTCCCAACACCTTTTGGTGAATGATGAAGACG	900
Qy	964	CCCTTCTATGATGCGCAAGCAACCTTAAAGGATTTTGTCTATGATGCTTTGGTCAATATAGTT	1023
Db	901	CCCTTCTATGATGCGCAAGCAACCTTAAAGGATTTTGTCTATGATGCTTTGGTCAATATAGTT	960
Qy	1024	GGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAGCTGTGAAGAAAT	1083
Db	961	GGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTCTGAAGCTGTGAAGAAAT	1020
Qy	1084	TGTAAGCCTAGAGTTCACCTGCGCATCTGCTTTTGAAGGACATATGTTACTCTGCTGCTA	1143
Db	1021	TGTAAGCCTAGAGTTCACCTGCGCATCTGCTTTTGAAGGACATATGTTACTCTGCTGCTA	1080
Qy	1144	GAGCCCTTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAAATGTT	1203
Db	1081	GAGCCCTTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAAATGTT	1140
Qy	1204	GCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGCAGGAACATATGAAGAACGCTTGTGT	1263
Db	1141	GCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGCAGGAACATATGAAGAACGCTTGTGT	1200
Qy	1264	GTTGCCAAGTGCAGGTGGAAATGGGAGCGCCAGGTGTTGGATGTCAACATGATGATGGC	1323
Db	1201	GTTGCCAAGTGCAGGTGGAAATGGGAGCGCCAGGTGTTGGATGTCAACATGATGATGGC	1260
Qy	1324	ATCCTAGATGCTTCAAGTGCATATGACAGATTTTGAACATTTAATTTGCTCCGAGCCAGAC	1383
Db	1261	ATCCTAGATGCTTCAAGTGCATATGACAGATTTTGAACATTTAATTTGCTCCGAGCCAGAC	1320
Qy	1384	ATCGAAAAGTACCTTTTGTGTCATGCTCTCCAAATTTTGTCTGTGATTTGAAGCTGGGTTA	1443
Db	1321	ATCGAAAAGTACCTTTTGTGTCATGCTCTCCAAATTTTGTCTGTGATTTGAAGCTGGGTTA	1380
Qy	1444	AAGTGTGCCAAGGGAAGTGCATTTGCAATAGCATTTAGTCTGAAGGAAGGAGGACGAC	1503
Db	1381	AAGTGTGCCAAGGGAAGTGCATTTGCAATAGCATTTAGTCTGAAGGAAGGAGGACGAC	1440
Qy	1504	TTCTTGGAGAGGCGCAGGAGATTTAAAGTATGGAGCTGCTATGTTGCTCATGGCTTTT	1563
Db	1441	TTCTTGGAGAGGCGCAGGAGATTTAAAGTATGGAGCTGCTATGTTGCTCATGGCTTTT	1500
Qy	1564	GATGAAGAAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGCTAC	1623
Db	1501	GATGAAGAAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGCTAC	1560
Qy	1624	CATCTGCTTGTGAAAAAATCTGGGCTTTTAAATCCAAATGACATTTATTTTGAACCTTAATC	1683
Db	1561	CATCTGCTTGTGAAAAAATCTGGGCTTTTAAATCCAAATGACATTTATTTTGAACCTTAATC	1620
Qy	1684	CTAACCATTTGGGACTGGAAATGGAGGAAACACAACTTGTATGCCATTAATTTTATCCATGCA	1743
Db	1621	CTAACCATTTGGGACTGGAAATGGAGGAAACACAACTTGTATGCCATTAATTTTATCCATGCA	1680
Qy	1744	ACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGCTCTTTTCAACTTG	1803
Db	1681	ACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGCTCTTTTCAACTTG	1740
Qy	1804	TCCTTCTCTTCCGAGGAATGGAGGCAATTCGAGGAAGCAATGATGGGGTTCCTTCTTAC	1863

Db 1741 TCCCTTCCCTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTCCTTTAC 1800  
Qy 1864 CATCAATCAAGTCTGGCATGACATAGAGATAGTAATGCTGGAACCTCCCTGTGTAT 1923  
Db 1801 CATCAATCAAGTCTGGCATGGAATGGGATAGTAATGCTGGAACCTCCCTGTGTAT 1860  
Qy 1924 GATGATATCCATATAGGAATCTTCGAGCTCTGTGAAAGATCTCATCTGGAATTAAGACCT 1983  
Db 1861 GATGATATCCATATAGGAATCTTCGAGCTCTGTGAAAGATCTCATCTGGAATTAAGACCT 1920  
Qy 1984 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGATC 2043  
Db 1921 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGATC 1980  
Qy 2044 ATTCAGACTGATAGTGAGAAATGGCCCTCTCGAAGAACCTTTGAGTATGCCCTTTGTG 2103  
Db 1981 ATTCAGACTGATAGTGAGAAATGGCCCTCTCGAAGAACCTTTGAGTATGCCCTTTGTG 2040  
Qy 2104 AAGGGCATTTGAAACATATTTATTTAGGATATCTGAGGAAGCCAGGTTAAACCAAAAAA 2163  
Db 2041 AAGGGCATTTGAAACATATTTATTTAGGATATCTGAGGAAGCCAGGTTAAACCAAAAAA 2100  
Qy 2164 TATCCCGACCTCTCAATATATTTAGAGGACCCCTGATGAATGAATGAATTTGGT 2223  
Db 2101 TATCCCGACCTCTCAATATATTTAGAGGACCCCTGATGAATGAATGAATTTGGT 2160  
Qy 2224 GATCTTTTGGAGCTGGAAATGTTTCTACTCAGGTTATAAAGTCAAGCCCGGTTATG 2283  
Db 2161 GATCTTTTGGAGCTGGAAATGTTTCTACTCAGGTTATAAAGTCAAGCCCGGTTATG 2220  
Qy 2284 AAGAGGCTGTGGCCACCTTATCCCTTTTCATGAAAGAAAGAGAAACAGAGTG 2343  
Db 2221 AAGAGGCTGTGGCCACCTTATCCCTTTTCATGAAAGAAAGAGAAACAGAGTG 2280  
Qy 2344 CTTAAAGGCAAGTAGAAGAGAGAGACCTTTACAGGGCACATCGTCTGGCCACTGTT 2403  
Db 2281 CTTAAAGGCAAGTAGAAGAGAGAGACCTTTACAGGGCACATCGTCTGGCCACTGTT 2340  
Qy 2404 AAAGCGAGCTGCAGCATAGGCAAGAAATAGTTGGAGTAGTCCCTGGCTGCAATAT 2463  
Db 2341 AAAGCGAGCTGCAGCATAGGCAAGAAATAGTTGGAGTAGTCCCTGGCTGCAATAT 2400  
Qy 2464 TTCCGAGTTATTTAGTAGTCACTGCTCATGATGATGATGATGATGATGATGATGATGAT 2523  
Db 2401 TTCCGAGTTATTTAGTAGTCACTGCTCATGATGATGATGATGATGATGATGATGATGAT 2460  
Qy 2524 GACCAAGAGAGATATATTTGGCTGTCAGAGACTCATCTCTCCCTGGATGAATG 2583  
Db 2461 GACCAAGAGAGATATATTTGGCTGTCAGAGACTCATCTCTCCCTGGATGAATG 2520  
Qy 2584 ATTTTGTTCAGAGGAATGGAGATTTAGCTATAGGATTTCCATTTGATTTGGAGGA 2643  
Db 2521 ATTTTGTTCAGAGGAATGGAGATTTAGCTATAGGATTTCCATTTGATTTGGAGGA 2580  
Qy 2644 GCAACCACTTCAAAAACCCACAGCAGTAAAAATAGCTCCGAGATACAGTGCACCTGTA 2703  
Db 2581 GCAACCACTTCAAAAACCCACAGCAGTAAAAATAGCTCCGAGATACAGTGCACCTGTA 2640  
Qy 2704 ATCCATGCTTCGAGCGCTCAGAGTGTGTGTGTTCCAGCTGTTAGATGAAT 2763  
Db 2641 ATCCATGCTTCGAGCGCTCAGAGTGTGTGTGTTCCAGCTGTTAGATGAAT 2700  
Qy 2764 CTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAGATTTAGCAGGACCAT 2823  
Db 2701 CTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAGATTTAGCAGGACCAT 2760  
Qy 2824 TATGAGTCTCTCAAGGAGGAGATATCTTACCTTTAAGTCAAGCCAGAAAAAGTGTTC 2883  
Db 2761 TATGAGTCTCTCAAGGAGGAGATATCTTACCTTTAAGTCAAGCCAGAAAAAGTGTTC 2820  
Qy 2884 CAAATGGATTTGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGTCTTTATTTGGACCCAGGTC 2943  
Db 2821 CAAATGGATTTGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGTCTTTATTTGGACCCAGGTC 2880

Qy 2944 TTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTCACCTGGAAGCCCTTTCTTTGAT 3003  
Db 2881 TTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTCACCTGGAAGCCCTTTCTTTGAT 2940  
Qy 3004 GTCTGGCAGCTCCGGGCAAGTACCAGATTCGAGGCTTCCCAAGATATTTAAACGACAAA 3063  
Db 2941 GTCTGGCAGCTCCGGGCAAGTACCAGATTCGAGGCTTCCCAAGATATTTAAACGACAAA 3000  
Qy 3064 ACAGTAGTGCAGAGGCCAGGAAGCTACGATGATGCCCAATATGCTGGAACACACTG 3123  
Db 3001 ACAGTAGTGCAGAGGCCAGGAAGCTACGATGATGCCCAATATGCTGGAACACACTG 3060  
Qy 3124 ATTAGTCAAAAGAAAATCCCGGCGCCGGGGTGTGGTTCCTGGCCAGACAGAGTATC 3183  
Db 3061 ATTAGTCAAAAGAAAATCCCGGCGCCGGGGTGTGGTTCCTGGCCAGACAGAGTATC 3120  
Qy 3184 CAAGACGACATTCACCTGTACGAGAGGCTGTGTGTGCCCCAGAGCTGAGAGCCCATAGCC 3243  
Db 3121 CAAGACGACATTCACCTGTACGAGAGGCTGTGTGTGCCCCAGAGCTGAGAGCCCATAGCC 3180  
Qy 3244 ACTTCTATGGTTAAGGCAACAGGCTGAGAAGACTCTGCCAGACGGAGCCATCTAC 3303  
Db 3181 ACTTCTATGGTTAAGGCAACAGGCTGAGAAGACTCTGCCAGACGGAGCCATCTAC 3240  
Qy 3304 TGCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCTGT 3363  
Db 3241 TGCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCTGT 3300  
Qy 3364 GCCGTTCCCTGCTTTGGGGTAGAAGACTGAGCAAGGCCCTATGAGGATGATGTGACGAC 3423  
Db 3301 GCCGTTCCCTGCTTTGGGGTAGAAGACTGAGCAAGGCCCTATGAGGATGATGTGACGAC 3360  
Qy 3424 TACAGCAGCATCATGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTGCAAGAGAG 3483  
Db 3361 TACAGCAGCATCATGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTGCAAGAGAG 3420  
Qy 3484 CTCCATCAAAAGAGTTCCGCCAGAACTGTGGCCCTACTGTGGCAGTGAAGCTGGAAGTC 3543  
Db 3421 CTCCATCAAAAGAGTTCCGCCAGAACTGTGGCCCTACTGTGGCAGTGAAGCTGGAAGTC 3480  
Qy 3544 GCAGACTGCGAAGTTTCCGCTACAAGGGCATCCGCCCGGCTCTGTGCTACCCAGCCAG 3603  
Db 3481 GCAGACTGCGAAGTTTCCGCTACAAGGGCATCCGCCCGGCTCTGTGCTACCCAGCCAG 3540  
Qy 3604 CCGACACACCCGAGAGCTCACCATGTGGAGCTCGCAGACATCGAGCAGTCTACAGGC 3663  
Db 3541 CCGACACACCCGAGAGCTCACCATGTGGAGCTCGCAGACATCGAGCAGTCTACAGGC 3600  
Qy 3664 ATTAGGTTAAAGTCAAAATATTTAGCAATGGCACTGCTTTCAGCAGTCTCAGGCTCTACTTC 3723  
Db 3601 ATTAGGTTAAAGTCAAAATATTTAGCAATGGCACTGCTTTCAGCAGTCTCAGGCTCTACTTC 3660  
Qy 3724 TCCAAATTTGAAGTCAAAATATTTTGTGTGGGGAGATTTCCAAAGATTCAGGTTGAGGAT 3783  
Db 3661 TCCAAATTTGAAGTCAAAATATTTTGTGTGGGGAGATTTCCAAAGATTCAGGTTGAGGAT 3720  
Qy 3784 TATGATTTAGGAGAAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTTTG 3843  
Db 3721 TATGATTTAGGAGAAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTTTG 3780  
Qy 3844 GGATATGATACAGACTAA 3861  
Db 3781 GGATATGATACAGACTAA 3798

RESULT 7

ADI39113  
ID ADI39113 standard; DNA; 3780 BP.

XX AC ADI39113;

XX AC ADI39113;

DT 22-APR-2004 (first entry)





Db 1286 TCGAGTTCTCTGAACTCATGCGCGCGAGCTGACATTTGCCCGTGTGCCGTCATGATCG 1345  
Qy 1409 ACTCTCTCAATTTTGTGTGATTAAGTCTGGTTAAAGTGTCTGCAAGGAGTGTGATTTG 1468  
Db 1346 ACTCATCCAAGTTCGAGATCATCGAGCGCGCTGNAATGCGTGCAGCGCAATTCGATCG 1405  
Qy 1469 TCAATAGCATTTAGTCTGAAGGAAGGAGGAGGAGCTTTCTTGGAGAAGCCAGGAAGATTA 1528  
Db 1406 TCAATTCATTTCCGTGAAGGAAGGCGAGGAGAAGTTTCTCCAGCAGGCTCGGCTCGTCC 1465  
Qy 1529 AAAAGTAGGAGCTGCTATGTTGGTCTATGCTTTTGAAGAGAAGGACAGCAACAGAAA 1588  
Db 1466 ACAATTACGGTGTGCGCGTGTCTGCTATGCTTGTGATGAGTGTGCGGCGAGCGGATACCT 1525  
Qy 1589 CAGACACAAAATCAGAGTGTGACCCCGGCTACCATCTGCTTGTGAAGAAAATCGGGCT 1648  
Db 1526 ATCAGCGCAAGGTGGAATCTGCGCGCGCTTCAAGCTTCTGACCGAAAGCCCGTTC 1585  
Qy 1649 TTAATCCAAATGACATTAATTTTACCTTAATATCCTAAACCATTTGGGACTTGAATGAGG 1708  
Db 1586 TGTCTCCGGAAGACATCATCTTTCGACCGGAATGTTTTCGCTGAGCTACGGGCAATCGAGG 1645  
Qy 1709 AACACAACTTGTATGCCATTAATTTTATCCATGCAACAAAGTCAATTAAGAAACATTTAC 1768  
Db 1646 AGCACATAATTAACGGGTGACTTCATCGAGGCCACCAAGACCATCGCGAATACCATGC 1705  
Qy 1769 CTGGAGCCAGAAATTAAGTGGAGGTCTTTCCAATTTGCTTCTTCTTCCGAGGAATGGAAG 1828  
Db 1706 CGCTCAGCGCATATTTCCGGGGGCGTTTCCAACTGTCTCTTCTTCCGCGGCAATGAGC 1765  
Qy 1829 CCATTCGAGAGCAATGATCGGGTCTTCTTTTACCAATCAAGTCTGGCATGGACA 1888  
Db 1766 CGGTGCGTGAAGGAGATGATCGGTGTTCTCTATCAGCCATTCAGTCTGGCATGGATA 1825  
Qy 1889 TGGAGATAGTGAATGCTGGAACCTTCCCTGTGTATGATGATATCCATAAGGAACTTCTGC 1948  
Db 1826 TGGGCATCGTCAACCGCGGCGAGCTGCGGTTTACGCAATATCGATGCGGAATGCGCG 1885  
Qy 1949 AGCTCTGTGAAGATCTCATCTGGAATTAAGACCTCTGAGGCCACTGAGAGCTCTTACGTT 2008  
Db 1886 AGGCTCGGAAGAGCTGTGTGCTGAACCCCGCGAGCATGCCACGAGCGTCTGCTCGAGG 1945  
Qy 2009 ATGCCAGACT- --CAAGGCAAGGAGGAGGAAGTCA- --TTCAGACTGATGAGTGA 2062  
Db 1946 TGGCGGAGCGTTTCCGTGTACGGGTGAAACAGGCCAAGGTGAGGATCTTTCCTGTC 2005  
Qy 2063 GAAATGGCCCTGTGGAAGAGCCCTTGAGTATGCCCTTGTGAAGGCAATTGAAAGAAACATA 2122  
Db 2006 CCGAGTATCCGTTGAAAGCGGCTGGAAATGCTCTGCTCAACGGCAATTACCGACTATA 2065  
Qy 2123 TTAATTGAGGATATGAGGAAGCCAGGTGTAACCAAAATAATATCCCGCACTCTCAATA 2182  
Db 2066 TCGAGGCCGATACGGAAGAGG- --ACGCCAGCAGCGCGCGCGCTGCTGATG 2116  
Qy 2183 TAAATTGAGGACCTTGATGATGGAATGAAATTTGTTGGTGAATCTTTTGGAGCTGAA 2242  
Db 2117 TCATCGAAGGCGCGCTGATGCGCGGTATGAATGTGGTGGTGACCTGTTCGGTTCCGGCA 2176  
Qy 2243 AAATGTTTCTACCTCAGGTATTAAGTCAAGCCCGGTTATGAAGAGGCTGTGTGCCACC 2302  
Db 2177 AGATGTTCTGCGACAGTGTGTGAATCCGCCGCTGTGATGAGCAGCGGCTGCGGTTTC 2236  
Qy 2303 TTATCCCTTTTATGGAAGGAGGAGGAGAAACAGAGTGTCTTAAACCGCAACAGTGAAG 2362  
Db 2237 TGTGCTCTTACATGGAAGAGGAA- --AGCGCCTGAATGCGCGTT 2278  
Qy 2363 AAGAGGACCTTACAGAGGACCATCTGCTGGCCACTGTTTAAAGGCACTGCAAGCA 2422  
Db 2279 CCGAGCGAGTGCCTCCCGGCAAGGTGCTAAATGCGGACCGGTGAAGGCGGACGTGCAAGATA 2338  
Qy 2423 TAGGCAAGAACATAGTTTGGAGTAGTCTTGGCTGCAATAATTTCCGAGTTATTTAGT 2482

Db 2339 TCGCAGAAACATCTGTGCGGCTTGTGTAGCTTGCACCAATTTACGAGATCATTTGATCTCG 2398  
Qy 2483 GAGTCATGACTCCATGTTGATAAGACTGAAAGCTGCTTTGACCAACAAGAGAGATATAA 2542  
Db 2399 GCGTGTGTTGTCGACGACGAAATCTCGAAACGGGATCGCGGAANAAGTGTGATGTGA 2458  
Qy 2543 TTGGCCCTGTGAGGACTCATCTCTTCCCTGGATGAAATGATTTTGTGTGCCAAGGAAA 2602  
Db 2459 TCGGCTCTCCGCTCATCACTCCGCTGCTGATGAGATGGTGTGATGTGCGCGCGGCA 2518  
Qy 2603 TGGAGAGATTAGCTATAAGGATTTCCATTTGATTTGAGGAGGAGCAACACTTCAAAAACCC 2662  
Db 2519 TGGAGCGACAGGGTTTCGACATTTCCGCTGCTGATCGCGGTGCGACACGAGCGTGTGC 2578  
Qy 2663 ACACAGCAGTTAAATAGCTCCGAGATAC- --AGTCACCTGTAAATCCATGCTCTGAGC 2719  
Db 2579 ATACGGCGTAAATAATCCATCCGCTTACGAGCGGCGAGGCGATCTATGTCACGACG 2638  
Qy 2720 CGTCCAAGAGTGTGTGTGTGTCTCCAGCTGTTTAGATGAAAATCTAAAGGATGAATACT 2779  
Db 2639 CTTGCGCGCGGTGGCGTGTCTTTCAGCGCTCTCTCCGAAGAGCAGAGCCGCTTATA 2698  
Qy 2780 TTGAGGAAATCATGGAAGAAATATGAGATATGAGACGACCATTTATGATCTCTCAAG 2839  
Db 2699 TCGACGGCATCCGAGCGGAATATGCCAAGGTGGCGGAAGCCATGCCCGCAATGAGCGCG 2758  
Qy 2840 AGAGGAGATCTTACCCTTAAGTCAAGCCAGGAAGAGTGTGTTCCTCAATCGATTTGGCTGT 2899  
Db 2759 AAAAGCAGCGCTTCCGCTTTCGCGCGCGGAGAAATGCGCAAGATCGACTG- --GT 2815  
Qy 2900 CTGAACCTCACCCAGTGAAGCCCACTTTATTTGGGACCCAGGCTTTTGAAGACTATGACC 2959  
Db 2816 CGAGTACAGGTTTTCAGCGCGAGTCTTTCGCGCAACAGGTTTTCGAGACCTATGATC 2875  
Qy 2960 TGCAGAAAGCTGTGTGACTACATTTGAGTGAAGCTTTTTCCTTGTGATGTCTGGAGCTCCGG 3019  
Db 2876 TGGAAAGCTTTCCGTTATCGACTTGAAGCGCGCTTCTTCCAGACCTGGGAAATGAAG 2935  
Qy 3020 GCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAAGCAACAAAGTGAAGTGAAGG 3079  
Db 2936 GCC- --GTTTCCGCGGCTTCTTGAAGACGAAACAGCGCGGAGCGG 2980  
Qy 3080 CCAGAAAGTCTACGATGATGCCCACTATGCTGAAACACACTGATTTAGTCAAAAGAAAC 3139  
Db 2981 CGCGGAGCTTTATGCGCATGCGCAGGCAATGCTTTCGAGAGATCATCGAGGAAGTGT 3040  
Qy 3140 TCGCGGCGCGGCTGTGTGCTGCGCAGCAGAGATATCCAAAGACATTCACCTCAC 3199  
Db 3041 TCCAGCACGCGGCTGATCGGCTTCTGCGCGGCAATGCGGTGGTGAAGATATCAGGC 3100  
Qy 3200 TGTACGAGAGGCTGTGTGCTGCCCGAGCTGCAAGCCCATAGCCACTTCTTATGGGTAA 3259  
Db 3101 TCTTTACGGAATGAAGTTCG- --AAGGAAGTGTGGGACGCTTCTTACGCTGC 3151  
Qy 3260 GGCAACAGGCTGAGAAGGCTCTGCCAGCAGCGGACCATACTACTGCTCTCTCAGACTTCA 3319  
Db 3152 GCCAGAGCTTTCAGAGCGGATGCGG- --TCCGAACGTCGCGCTGTCCGATTCG 3205  
Qy 3320 TCGCTCCCTTGTCAATCTGGCATCCGTCACTACCTGGGCTGTGTTCGCT- --TGCCTGCT 3376  
Db 3206 TCGCGCCGCTCGAATGCGGCTTTCGCGGCTTATGCTGCGGCTTTCGTGTAAACGCGGTA 3265  
Qy 3377 TTGGGTTAGAAGAGCTGAGCAAGGCTATGAGGATGATGTGAGCATCTACAGCAGCATCA 3436  
Db 3266 TCGAGGAAGTGGCGATTTGCCGAGCGCTTCGAGCGGCGCAATGACGATTTATTCGTCCTCC 3325  
Qy 3437 TGGTCAAGGCGCTGGGAGCCGCTGCGAGAGGCTTTGCAAGAGAGCTTCCATGAAGAG 3496  
Db 3326 TCGTCAAGGCTTGTGATGACGCTTTTCCGAGAGCTTTCGCGAGGCTATGATGAGCGG 3385  
Qy 3497 TTCGCCGAGAACTGTGGGCTTCTGTGCGAGTGAAGCTGGAAGTGTGCGAGACCTCGGAA 3556  
Db 3386 TGCCAAGGAGTCTCGGGTTATGCGCGGAGGAGGCTTTCGCGGTGACGATCTGATAG 3445

QY 3557 GGTTCGGGTACAGGGATCCGCGGCTCTGCTACCCAGCCAGCCGACACACCG 3616  
 DB |||||  
 DB 3446 GCGAAGCTTATGCGGTATCCGCGGACCGGTTATCCGGCCAGCCGACACACCG 3505  
 QY 3617 AGAAGCTCACCATCTGGAGACTCCGACACATCGAGCAGTCTACAGGCATTAGTTACAG 3676  
 DB |||||  
 DB 3506 AAGAAGACGCTGTTGCTCTGCTGAGCCGCCAATGCGCGGTGTGGAATTGACGG 3565  
 QY 3677 AATCATTTAGCAATGGCACCTGCTTTACAGCAGTCTCAGGCCCTCTACTTCTCCAATTGGAAGT 3736  
 DB |||||  
 DB 3566 AAGCTATGCGATGTGCGCGGCTCGTCTGCTGAGCCGCCAATGCGCGGTGTGGAATTGACGG 3625  
 QY 3737 CCAATATTTGCTCTGGGAGAGATTTCCAAAGATCAGTTGAGGATATGCAATTGAGGA 3796  
 DB |||||  
 DB 3626 GCTATTATTTGCGGCTTGCACAGGTGAGCGGATCAGGTTCTGACTATGCGCGCGCA 3685  
 QY 3797 AGACATATCTGCTGCTGAGTCTCAGAAATGGCTTGGACC 3836  
 DB |||||  
 DB 3686 AGATATGCGGTACAGAGGTGAGCGCTGGCTCGCGCC 3725

## RESULT 8

ACA37652  
 ID ACA37652 standard; DNA; 3786 BP.

AC ACA37652;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #19309.

XX Antisense; db; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.

XX Mycobacterium avium.

XX W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU33782.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 25522; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 3786 BP; 695 A; 1293 C; 1265 G; 533 T; 0 U; 0 Other;

Query Match 23.3%; Score 914.6; DB 8; Length 3786;

Best Local Similarity 54.8%; Pred. No. 2.7e-249;

Matches 2055; Conservative 0; Mismatches 1634; Indels 60; Gaps 10;

QY 121 GATGAGATCAATGCCATTCTGACAGAGAGGATTATGCTGATGAGGAGGATGGGACC 180  
 DB |||||  
 DB 82 GAGCACTGACGGCGCGCTGCGCCAGCGGATCATGTTGATCGAGCGGCATGGGACC 141  
 QY 181 ATGATCCAGCGGAGAGCTTAAACGAAGAACATTCGAGGTGAGGAATTTAAAGATCAT 240  
 DB |||||  
 DB 142 GCGATCCAGCGGAGCGCCCGCCAGCGCGGCTACCGCGCGAGCGGTTCAAGGACTGG 201  
 QY 241 GCCAGCGCGCTGAAGGCAACATGACATTTTAAAGTATTAATCACTCAGCTGATGCTTAC 300  
 DB |||||  
 DB 202 CCGAGCGATCTGTCGGCAACACGACCTGCTCAACTGACGAGCGCGAGATCGAG 261  
 QY 301 CAAATCCATAAGGAATACTTGTGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGC 360  
 DB |||||  
 DB 262 GGCATTCACCGCGAGTACCTCGAGCGCGCGCGACATCTCGAGACCAACACCTTCAAC 321  
 QY 361 AGCACTAGTATTGCCAAGCTGACTATGGCTTTGAAACACTTGGCCTACCGGATGAACATG 420  
 DB |||||  
 DB 322 GCGAAGCGCATCTCGCTGTCGACTACGGCATGGAGAGCTGAGTACGAGCTGAACATAC 381  
 QY 421 TGCTCTGAGGAGTGGCGAGAAAAGTCCGAGGAGTAATACTCTCCAGACAGGAATTAAG 480  
 DB |||||  
 DB 382 GCGGCGCGCGCTGCGCGCGTAAGCGCGCGAGGTTACAGCACCC---CGGACAAAGCCC 438  
 QY 481 AGGTTTGTGCGAGGGCTCGGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTG 540  
 DB |||||  
 DB 439 CGCTACGTGCGCGCGCGCTGCGCGCGAGCTCGCGGACCGCGTGTCTCGCGGACGTC 498  
 QY 541 GAAAGGCGCGATTATAGGAACATCACATTTGATGAGCTTTGTAAGCATACCAAGACAG 600  
 DB |||||  
 DB 499 AACGATCCCGGGCGCGCAACGTCACCTACGAGAGCTGTGTCGCGCTACCTCGAGCGG 558  
 QY 601 GCCAAGGACTTCTGGATGCGGGTTGATATCTTACTTCAATGAAACTATTTTGTACT 660  
 DB |||||  
 DB 559 GCCAAGCGCTGTGTCGACGCGGTTGCGACCTCTGATCTGTCGAGAGCATCTTCGACAG 618  
 QY 661 GCCAATGCCAAGCAGCTTGTGCTTCACTTCAAAATCTTTTTCAGGAGAAATATGCTCCC 720  
 DB |||||  
 DB 619 TTGAAGCCCAAGGCTCGGTTGTCGCCCTCGAGACGCTGTTCGAACAGCGCGCGCGCG 678  
 QY 721 CGGCTTATCTTTATTTTCAGGAGCATGCTTTGTAAGTGGCGGACTCTTTTCGGACAG 780  
 DB |||||  
 DB 679 TGCGCGGTATCATCTCGGCGCAACATCAGCATGCTCCGCGGAGCGCTGTCCGGGCG 738  
 QY 781 ACAGGAGAGGGATTTGTATCATCGGTGCTCTCATGGAGAACCACTCTGTCATTTGAAT 840

Db 739 GTCCAGGCGTTCTGGAACTCGATCGGCAAGCCGATCGCGTCTGAGCTCAAC 798  
Qy 841 TGTGCTTGGGTGAGCTGAGATGAGACCTTTTATTGAAATAATTGGAAATGTACAAACA 900  
Db 799 TGCGCCCTGGGCGACCGGAGATCGGCCCTTACATCGCCGAGGTGTGCGGCATCGCGGAC 858  
Qy 901 GCCTATGTCCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTGGTGTGACTATGATGAA 960  
Db 859 ACCTTGTCTCTGCTACCCGAAAGCCGCGCTTGCCCAACGCTTCGCGGAGTACGACGAG 918  
Qy 961 ACGCCCTTCTATGATGGCCCAACACCTAAAGATTGTTGCTATGATGGCTTGGTCAATATA 1020  
Db 919 TCCCGGAGCGTCAAGCCTCTACATCGCGACTTCGCGAGCGCGCTTGTCAACATC 978  
Qy 1021 GTTGAGAGTCTGTGGGTCAACACGAGATCATATCAGGGAAATGTGTAAGCTGTGAAA 1080  
Db 979 GTCGCGGTGCTGCGGAACCGCGCGCGCACATCGCCGAGATCGCCAAAGTGTGCGAG 1038  
Qy 1081 AATTGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGGT 1140  
Db 1039 GGCAAAGCGCGCG-----CGAGGTGCGCAGATTTCGGGTGCCACCCG--GCTCTCGGBC 1092  
Qy 1141 CTAGAGCCCTTCAGGATTGGACCGGTACACCAACTTTTGTAAATTTGGAGAGCGCTGTAAT 1200  
Db 1093 CTCGAGCGCTCAACATCACCGAGCTCGCTGTTGTGTAACATCGCGGAGCGCACCAAC 1152  
Qy 1201 GTTGAGAGTCAAGGAAGTTTGTCTAACTCATGTGGCAGGAAACTATGAAGAAGCCTTG 1260  
Db 1153 ATCACCGGCTCGGCCGTTCCGCAACCTGATCAAGGCCGAGGACTACGACACCGCGCTG 1212  
Qy 1261 TGTGTGCCAAAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCAACATGATGAT 1320  
Db 1213 TCGGTGCGCTGTGACGAGTGAAGTCTGTCGCCAGGTTCATCGACATCAACATGACGAG 1272  
Qy 1321 GGCATGTCTAGATGGTCAAGTGCATGACAGATTTTGGCAACTTAAATGCTTCGAGGCA 1380  
Db 1273 GGCATGATCGACGGGTGCGCGGATGGACGGTTCCAGGCTGATCGCGCGGAGCG 1332  
Qy 1381 GACATCGCAAGGTACCTTTGTGATCGATCTCTCCCAATTTTGTGTGATGAAGCTGGG 1440  
Db 1333 GACATCAGCGGTTCCCGGTGATGATCGACTCTCTCAAGTGGAGGTCAICGAGCGCGC 1392  
Qy 1441 TTAAGTCTGCCAAGGAAGTGCATTTGCTCAATAGCATTAGTCTCAAGGAGGAGGAC 1500  
Db 1393 CTGAAGAACGTGACGGGAAGCCGATCTCACTGATCTCTTTGAAGAGGGCGAGGAG 1452  
Qy 1501 GACTTCTTGGAGAAGCGCAGGAAGATTAAAGATGAGGCTGTATGTTGTTGCTCATGGCT 1560  
Db 1453 AAGTTCTGTCGCGAGGCGCGCTGTGCGCAAGTACGGCGCGCGTCTGTTGATGGCC 1512  
Qy 1561 TTTGATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGATGTGCAACCGCGGCC 1620  
Db 1513 TTCGACGAACAGGGTCAGGCGCGACAACTGGAGCGCGCCAAACAGATCTGCGCACGCGCC 1572  
Qy 1621 TACCATCTGCTGTGAATAAAGCTGGCTTTAATCCAAATGACATTTATTTTGAACCTAAT 1680  
Db 1573 TACCCTGCTGTGACCGGAAGAGTCTGGCTTCGCGCCGAGGACATCATCTTCGACCCCAAC 1632  
Qy 1681 ATCTCTAAACATTTGGAGCTTGAATGAGGAAACACAACTTGTATGCGCATTAATTTTATCCAT 1740  
Db 1633 TGCCTCGCGTGGCGACCGGCATCGAGGACACGCAACTACGGCATCGACTTCATCGAG 1692  
Qy 1741 GCAACAAAGATCTTAAGAAACATTAATCTGGAGCCAGAAATGATGGAGGTCTTCCAAAC 1800  
Db 1693 GCCTCGCGCTGATCAAGAGAACTGCGCGGGGTGCACCTCTCGCGGGCATCTCGAAC 1752  
Qy 1801 TTGCTCTTCTCTCCGAGGAATGAAGCCTATCCAGAGCAATTCGATGGGTTTTCCTT 1860  
Db 1753 GTGTCGTTCTGTTCCGGGAAACAAACCCGTCGCGAGGCGATCCATCTCGGTGTTCTG 1812  
Qy 1861 TACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGAAATGTGGAACCTCCCTGTG 1920

Db 1813 TTCCAGCCATCAAGCGCGCTGGACATGGGCATGCTCAACGCGCGCGCTGGTGCCC 1872  
Qy 1921 TATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTCGAATAAAGAC 1980  
Db 1873 TAGACTCGATCGACCCGAGCTGCGGACCGCATCGAGGACGTCTGCTGTAACCGCGCG 1932  
Qy 1981 CCTGAGGCCACTGAGAAGCTCTTACGTTATGCCAG-----ACTCAAGGCACAGGAGGG 2034  
Db 1933 GAGGACGCGCGCGAGAGGCTGTGGAGATCGCCGAGCGGTTCAACAAGTCGCGAGATGCT 1992  
Qy 2035 AAGAAATGCTATTCAGACTGATGATGAGTGGAGAAATGGCCCTGTGGAAGAACCGCTTGAGTAT 2094  
Db 1993 TCGAGGATTTCCGCGCGCGCGAGTGGCGCGCTGCGCGTCCGCGAGCGGATCACGCAC 2052  
Qy 2095 GCCCTGTGAGGAGGCAATTGAAACAATATTTTGAAGATACTGAGGAAGCCAGG---TTA 2151  
Db 2053 GCCCTGTGTAAGGAGCATCGACGCCACGTCGAGGACACCGAGGAGTTGCGGCGCGAG 2112  
Qy 2152 AACCAAAAAAATATCCCGACCTCTCAATATATTAATGAAGAGACCCCTGATGAATGGAATG 2211  
Db 2113 ATTGCCCGCGCGGTGCGCGCGATCGAGGTGATCGAGGCGCGCTGATGGACGCGATG 2172  
Qy 2212 AAAATTGTTGTTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAAGTTATAAAGTCA 2271  
Db 2173 AAGTCTGTCGCGACCTGTTTCGCTCCGCAAGATGTTTTCGCCCAAGTGGTGAAGTCG 2232  
Qy 2272 GCCCGGTTATGAAGAGGCTGTGGCCACCTTATCCCTTTTCATTTGAAAAAAGAAAGAA 2331  
Db 2233 GCCCGGTTGTAAGAAAGGCGCTGCGCTACTCTGCTGCTTCATCGAGGCGGAGAAAGAA 2292  
Qy 2332 GAAACACAGAGTGTCTTACGGGCACAGTAGTAAGAAAGAGACCCCTTACCAAGGCGACCATCGTG 2391  
Db 2293 GAGTCCGCGGTGTCGGTTCCA-----AGGACACCAACGCGACCATCGTG 2337  
Qy 2392 CTGCGCACTGTTTAAAGCGGACGTGCGACGACATAGGCAAGAACATAGTTGGAGTAGTCCTT 2451  
Db 2338 ATGCGACCGTCAAGGCGGACGTCACGACATCGGCAAGAACATCGTGGAGTTGTCTTG 2397  
Qy 2452 GGCTGCAATTAATTCGAGTTATTGATTAGGAGTCATGACTCCATGCTGATAGATACTG 2511  
Db 2398 CAGTGCACAACTTCGAGGTGATCGACTCGGTGATGTTGGCGCGGCCACGAAGATCCTC 2457  
Qy 2512 AAAGCTCTCTTTGACCAAGCAGAGATATAATTTGGCCTGTGAGGACTCATCACTCTCTTC 2571  
Db 2458 GATCTCGGAAAGACGACGCGGACATCATCGGCTGTTCGCGCTGATCACCCCGTCC 2517  
Qy 2572 CTGATGAATGATTTTGTGCAAGGAAATGAGAGATTTAGCTATAAGGATTCCATG 2631  
Db 2518 CTGACGAGATGTCCAACCTTCGCGCGGAGATGGAAACGCGAGGCGCTGACAGATCCCGCTG 2577  
Qy 2632 TTGATTGGAGGACCAACCTTCAAAAAACCAACACAGCAGTTTAAATAGCTCCGAGATAC 2691  
Db 2578 CTGATCGCGCGCGGACCACTTCGCGCGCCACACGCGCCGTAAGAGTCTCGCGCGCGGT 2637  
Qy 2692 AGTCACCTGTAATCCATGCTCGGACGCTCCAAAGAGTGGTGGTGTTCCTCCAGCTG 2751  
Db 2638 TCCGCGCGGTGCTCTGGTCAAGACGCTCCCGCTCGGTGCGCGCTGCGCGCGCTG 2697  
Qy 2752 TTAGATGAATAATTAAGGATGAATACTTTTGAAGAAATCA TGGAAAGATATGAAGATAT 2811  
Db 2698 CTGACGACAGCAGCGCGCGCGCTCTGAGAGCCACCGAGAGAGGACTACGCGTCTGCT- 2756  
Qy 2812 AGACAGGACCAATATGATCTCTCAAGAGAGGAGATACTTACCTTTAAGTCAAGCCAGA 2871  
Db 2757 --GCGGAAACGGCAGCGCCAGAAACGAGCGCGGATGGTGA CGCTGGAAAAAGGCCCGC 2814  
Qy 2872 AAAAGTGTCTTCCAAATGGATTGGCTCTCTGAACCTCACCCAGTGAAGCCACCGTTATT 2931  
Db 2815 GCCAACCGGACCGCATCGACTGGGACGGCTACACGCGCGCGTGCCTCGCATCGGTGCG 2874  
Qy 2932 GGGACCCAGGCTTTTGAAGACTATGACCTGCAAGAGCTGTGGAGCTACATTTGACTGGAG 2991  
Db 2875 GGAATA CGGGAATTTTCAGGACTACGACCTCGCGAGCTGCGCGAGTACATCGACTGGCAG 2934

```
Qy 2992 CTTTCTTCTGATGCTCGAGCTCGGGGCAAGTACCCGATCGAGGCTTCCCAAGATA 3051
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 2995 CCGTCTTCAACGCTGGGAGT-----GAGGGTCTGCTTCCCCGCAATC 2979
Qy 3052 TTTAACGACAAACAGTAGTGGAGAGCCAGGAGGTCTACGATGATGCCCAATATG 3111
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 2980 CTCAACAACCGGCCACCGCGGAGCGCGCCGCAAGCTCTACGACGACGCCAGGAGATG 3039
Qy 3112 CTGAACACATGATAGTCAAAAGAACTCCGGGCCCGGGGTGGTGGTTCGGCCA 3171
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3040 CTCGACACCTTGATCAAGAGAGTGGCTGACGCGCCCAACGCGGTGATCGGGTTCCTCCG 3099
Qy 3172 GCACAGAGTATCCAAAGACGACATTCACCTGACGACGAGGCTGTGTGCCCGGCTGCA 3231
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3100 GCGAACGGATCGCGCGCGGTTTGAAGACATCGAGCTGTACACCGAGCACCCCGCAC 3159
Qy 3232 GAGCCCATAGCACTTTCTATGGGTAAAGGCAACAGCGCTGAGAGAGACTCTGCCAGCACG 3291
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3160 GAGTGTCTGACCACTGTCACAACTCTGCGCAGCGGCGGAGCACCGGACGGNATC--- 3216
Qy 3292 GAGCCATCTACTGCTCTCAGACTTCATGCTCCCTTTCGATCTTGGCATCCGCTGACTAC 3351
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3217 ---CGGAACCGTCTGCTGGGCACTAGCTGCCCGCCCAAGAAACGGGCGACCGGACTAC 3273
Qy 3352 CTGGGCTGTTTGGCGT---TGCCTGCTTTGGGTAGAGAGCTGAGCAAGGCTCTATGAG 3408
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3274 ATCGGCGCTTTCGGCGGTGACCGCGGGCTGGGCGAGCGAGAGATCGCGGAGTTCAAG 3333
Qy 3409 GATGATGTCGAGCTACAGCAGCATCATGCTCAAGGCGCTGGGGACCGGCTGGCAGAG 3468
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3334 GCGGCTTGGAGCTACAGCGCATCTCTGAGTGTGATCGCGGACCGGCTGGCGGAG 3393
Qy 3469 GCCTTTGCGAAGAGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGT 3528
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3394 GCGTTCCGCGAAGGATGACCAAGCGGTCCGCAAGGAGTTCTGGGGTACCGCCGAC 3453
Qy 3529 GAGCAGTGGAGCTGCGAGACTGCGAAGTTGCGGTACAAAGGCGATCCGCGCGCTCCT 3588
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3454 GAGCACTGGACAACGAGCACTCATCGACGAGAAGTACCGGGGAATCCGCGCGCGCG 3513
Qy 3589 GGCTACCCCGCAGCCGACCCAGCAGGAGCTCACCATGTGCGAGACTCGCAGACATC 3648
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3514 GGCTACCGGCTTCCCGGAGCACACCGAGAGGTGACGCTGTGGAAGTTGATGAGCTC 3573
Qy 3649 GAGCAGTCTACAGCATTAGGTTAAACAGATCATTAGCAATGGCACCTGCTTACGACATC 3708
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3574 AAGGAGCGCACCGGATCGAGCTACGGAGTGTGATGGCCATGTGCCCGCGCGCGCTC 3633
Qy 3709 TCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAG 3768
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3634 AGCGTTGGTATTCTCGACCCGCGAGTCGACGATCTTCGTGTGGGCGGTTGGCCAG 3693
Qy 3769 GATCAGTTGAGGATTTGATGAGGAGAACTATCTGTGGCTGAGGTTGAGAAATGG 3828
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3694 GACCAAGTTCGCGACTACGCGAGCGCAAGGCTGAGCTGGCGGCGGCGGCGGCTGG 3753
Qy 3829 CTTGAGCCCATTTGGGATGATACAGA 3857
Dd ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 3754 CTGCGCCCGCAACCTCGGCTACACCCCGA 3782
```

## RESULT 9

ACA20972

ID ACA20972 standard; DNA; 3684 BP.

XX AC

XX ACA20972;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #2629.

XX DE

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW

drug design; gene.

Acinetobacter baumannii.

WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

P-PSDB; ABU17102.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 8842; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 3684 BP; 1088 A; 747 C; 833 G; 1016 T; 0 U; 0 Other;

Query Match 23.3%; Score 911.8; DB 8; Length 3684;

Best Local Similarity 55.3%; Pred. No. 1.6e-248;

Matches 2070; Conservative 0; Mismatches 1557; Indels 114; Gaps 11;

Qy 128 TCAATGCCATTCTGCAGAGAGGATTTATGCTGCTGATGAGGATGGGACCATGATCC 187

Dd 20 TAAAGCGCTTCTGCTAAACGCACTTCTGATTTATGTTGTCATGGGACCATGATCC 79

Qy 188 AGCGGGAGAGCTAAACGAGAACACTTCCGAGTCAGGAATTTAAAGATCATGCCAGGC 247



Db 80 AGCCCAATAATTTGGAAGAAGCTGACTATATCGTGTGAGCGTTTGGCTGATTTGGSCACATG 139  
Qy 248 CGCTGAAAGGCAACAATGACATTTTAAAGTAACTCAGCCGTGATGTCATTTACCAATGCC 307  
Db 140 ACTTAAAGGTAACAATGACCTTTTGGTTCTTAAACACAGCCCTCAATCATTTCAAGGTATTC 199  
Qy 308 ATAAAGGAATCTGCTGGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCAGCACTA 367  
Db 200 ATGAAGCCTACTCGATGCTGGTGCAGACATATTTGAAACCAAGCTTTTAAACCGGTACAC 259  
Qy 368 GTATTGGCCCAAGCTGACTATGGCTTGAACACTTTGGCCTACCGATGAACATGTGCTCTG 427  
Db 260 GTGTTTCAATGCTCTGACTATCACAATGGAAGATCTGTTCCAGAGATTTAAACCGTGAAGCAG 319  
Qy 428 CAGAGTGGCCAGAAAAGCTGCCAGGAGGTAACTCTCCAGACAGGAATTAAGAGGTTTG 487  
Db 320 CACGTTTAGCTTAAAGCAGCTTGGCAAAAG---TATTCAACGCCACAGAACGCTCGTTTTG 376  
Qy 488 TGGCAGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGTCGCCCATCTGTGGAAGGC 547  
Db 377 TGGCAGGTGTACTTGGGCAACATCTCGTACATGTTCAATCTCTCCAGATGTGAACAACC 436  
Qy 548 CGGATTAAGAAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCCCAAAG 607  
Db 437 CTGCTTTTGTAAATTTAGCTTTGATGACTTAAAGAAATTTATTTGAAGCGACTCATG 496  
Qy 608 GACTTCTGGAATGGCGGGTTGATATCTTAATCTCATTTGAACACTATTTTGTATCTCCAAATG 667  
Db 497 CACTAAATGAAGTGGTGCAGACATTTCTCGATCGAAACTGTATTTGATACTTTAAACT 556  
Qy 668 CCAAGGAGCCTGTTTGCATCTCCAAATCTTTTGTGAGGAATATGCTCCCGGCTTA 727  
Db 557 GTAAAGCAGCGAATTTTGTGAGTCAAGAAAGCTTTTAAACAAATTTGGTGGCAATTAACAA 616  
Qy 728 TCTTTATTTTCAAGGACGATCTTGATAAAGGTGGCGGACTCTTTCGGGACAGACAGGAG 787  
Db 617 TTATGATTTTCAAGGACCAATTAACGATGATCAGGCCGTACTTTAAACAGGTGAGACAGCGG 676  
Qy 788 AGGATTTGTATCAGCGGTGCTCATGAGAGAACCACTCTGCAATTTGAATTAATTTGCTT 847  
Db 677 AAGCTTTCTGAACTCGTTCGTCATGCGGATTTGCTTTCAATCGGTTTAACTGTGCCC 736  
Qy 848 TGGGTGAGCTGAGTGAAGCTTTTATTTGAATTAATTTGGAATTTGGAATTTGACACGCTTATG 907  
Db 737 TTGTTGACAGATGCAATGCGCCCTCAACGTAATACTATTTCCGATGTGCGAGTACCTTTG 796  
Qy 908 TCCTCTGTTATCCCAATGCAAGTCTTCCAAACACTTTTGGTGAATGATGATAAGAGCGCTT 967  
Db 797 TTTCAAGCCCAACCAATGCAAGCTTACCAACGCAATTTGGTGAATATGACGAATCTCCAG 856  
Qy 968 CTATGATGGCCAGACACTAAAGGATTTGCTATGGAATGGCTTGGTCAATATATGTTGAG 1027  
Db 857 AGCAAACTGACGCTTTCTTAAAGAGTTTGTGAAAGCGGTTTGAATTAACATTAATCTGGTG 916  
Qy 1028 GATGCTGGGTCAACACCAATCATATCAGGGAATTTGCTGAAGCTGTGGAATTTGTA 1087  
Db 917 GTTGTGTTGGTACGACACCAACCATATTTGAGCTATTTGCCAATTTGCCGTAATAAGACATG 976  
Qy 1088 AGCTTAGAGTTCCACCTGCACTGCTTTTGAAGGACATATGTTACTGTCTGCTAGAGC 1147  
Db 977 CGCTTCGCAAGTGGTGAACCGTACCTGCTTGC-----GCTTAAGTGGTTTAGAAC 1030  
Qy 1148 CTTTCAGGATTTGGACCGTACCAACTTTGTTAACTTTGGAGAGCGCTGTAATGTTGCAG 1207  
Db 1031 CATTTAATATTTATGATGATTTTCAATTTGTTGTTAAAGCTTTGGGAGCGTACTTAACGTTACC 1090  
Qy 1208 GATCAGGAAGTTTGTCTTAACATCATCGGAGGAACATATGAAGACCTTTGTGTTG 1267  
Db 1091 GTTCTTAAATAATTTCTTACGCTCTCATCTGTAAGAAACCTTTTGCAGAAAGCTTTAGAAGTTG 1150  
Qy 1268 CCAAGTGCAGGTGGAAATGGAGCCAGGTTGATGTCAACATGATGATCATGGCATGC 1327

Db 1151 CACAGCAGCAGGTTGAAGCTGGCGCACAGATTATCGACATTAACATGATGAAGGATGC 1210  
Qy 1328 TAGATGTTCCAACTGCAATGACACAGATTTTGCAACTTAATTTGCTTCGAGCCGACACATCG 1387  
Db 1211 TCATTTGCAAAATGCGATGGTGCAATTTCTTAAACCTTTAGCATCGGAACCGGACATTT 1270  
Qy 1388 CAAAGGTACTCTTTGTGCATCGACTCCTCCCAATTTTGTCTGATTTGAAGCTGGGTAAAGT 1447  
Db 1271 CAGGTGACCGATCATGATTTGACTCATCGAATGGGAAATCATTTGAAGCCGGCTTAAAT 1330  
Qy 1448 GCTGCCAAGGAAAGTGCATTTGCTCAATAGCAATTAAGTCTGAAGGAAGGAGAGACGACTTCT 1507  
Db 1331 GCGTACAAGGTAAACCGGTTGTTAACTCAATTTCTTAAAGAAAGGTTATGACGAGTTG 1390  
Qy 1508 TGGAGAAGGCCAGGAAGATTTAAAGATATGAGCTGCTATGGTGGTCAATGCTTTTTCATG 1567  
Db 1391 TTGAAAGAGGCCCTCTGCGCTCAATATGCTGCTCAATTTATTTGTGATGGCTTTTGACG 1450  
Qy 1568 AAGAAGGACAGGCAACAGAAACAGACACAAATCAGAGTGTGCACCGGGCTTACCATC 1627  
Db 1451 AAGTAGTCAAGCCGACACATGCTGAGGTAACAGTGAATCTGTAGCGCTCTTATGACA 1510  
Qy 1628 TGCTTTGTAATAAATCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATATCCTAA 1687  
Db 1511 TTTTGGTTAACGAAGTAGGCTTCCCTGCTGAGATATTAATTTTGAACCGAACTGTTG 1570  
Qy 1688 CAAATGGAATCTGGAATGGAGGAAACAACTTTGTATGCAATTAATTTTATCCATGCAACAA 1747  
Db 1571 CAGTTGCGACTGCTGTTTGAAGAACATAACAACTACGCACTCGATTTTATTTGAAGCAACGG 1630  
Qy 1748 AAGTCATTAAGAAACATTAACCTGGAGGCAAGATAAGTGAGGTCTTTTCCAACTTTGCTCT 1807  
Db 1631 GCTGGATTAACAGAACTTACCGACGCCATGATTTCTGTTGTTGTTCTTAAGCTTTGCT 1690  
Qy 1808 TCTCTTTCCGAGGAATGGAAAGCCATTTGGAAGCAATGCAATGGGTTTCTCTTTTACCATG 1867  
Db 1691 TCTCATTTCCGTGGCAATGAACAGTTCTGAGGCCATTCCTCTGATTTCTTGTACCATG 1750  
Qy 1868 CAATCAAGTCTGGAATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTGTATGATG 1927  
Db 1751 CCATCAAGCAAGGATGACCATGGGTATTTGTGAACGAGGTCAATTTGGCTTATTTATGATG 1810  
Qy 1928 ATATCCATTAAGAACTTCTGCACTGCAAGATCTCATCTGGAATAAAGACCTCTGAGG 1987  
Db 1811 ATATCTCTACCGACTTAAAGAGCGTTGAGATGTCATTTTAAACCAAAATCAAGGTG 1870  
Qy 1988 CCACTGAGAAGCTTTTACGTTATGCCCCAGACTCAAGGACA----- 2028  
Db 1871 AGTCTGGTCAAGTCGCGACTGAGAAACTACTTGAAGTTGCAGAAAAATACCGTGGACAAG 1930  
Qy 2029 --GGAGGGAAGAAAGTCAATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCC 2086  
Db 1931 GTGGTGCAACAAAGGAAGCCGAGAACCTTGAATGGCGTAAATGAGTTCAGTTGAAATAACGTC 1990  
Qy 2087 TTCAGTATGCCCTTGTGAAGGCAATTTGAATAACATATTTGAGGATCTAGGAGAGCA 2146  
Db 1991 TTGAATATGCTTGTGTTAAAGGTTATACGACTTATATTGACCAAGACCCGGAAGAGGCC 2050  
Qy 2147 GGTTAACCAAAAAAATAATCCCGCACTCTCAATATAATTTGAAGGACCCCTGATGAATG 2206  
Db 2051 GCTTAATAATCAAA-----CGTCTTTTAGATGTAATTTGAAGGGCCACTGATGGACG 2101  
Qy 2207 GAATGAATAATTTGTTGGTGAATTTTGGAGCTGGAAAAATTTTCTACCTCAGGTTTAA 2266  
Db 2102 GCATGAACGTCGTCGGTGAATTTGTCGTTTCAAGCAAAATGTTCTTGCCCAAGTCTGTAA 2161  
Qy 2267 AGTCAGCCCGGTTTGAAGAGGCTTGTGGCCACCTTTATCCCTTTCTATGTAAGAAAGAA 2326  
Db 2162 AATTCGCCCGAGTCATGAACCAAGCAGTAGCATGGCTCAACCCGTCATCATGAGACTGAAA 2221  
Qy 2327 GAGAAGAAACCAAGAGTCTTTAAACGGCACAGTAGAAGGAAGAGGACCTTTACAGGGCACCA 2386  
Db 2222 -----AGACCGAAGGACAGTCTTAAAGGTAAG 2248



```
QY 2387 TCGTCTGGCCACTGTTAAAGCGGACGTGACGACATAGGCAAGAACATAGTTGGAGTAG 2446
Db 2249 TCCTAATGCGCAACCGGTTAAAGGTGACGTACACGATATTGGTAAATAATTGTAGCGGTAG 2308
QY 2447 TCCTTGGGTGCAATAAATTTCCGAGTTATTGATTTTAGGAGTCACTACTCCATGTCATAGA 2506
Db 2309 TACTTGGGTGTAATGGCTATGACATTTGTGACCTTGGCGTAAATGGTCCCTTGGGAGAAA 2368
QY 2507 TACTGAAAGTGTCTCTGACCAACAAGCAGATATAAATTTGGCTGTGACGACTCATCTC 2566
Db 2369 TCTTACAACTGCAATGATGAAATAATGTGACATCATTTGGACTGTCTGCTGATCACCC 2428
QY 2567 CTTCCCTGGATGAATGATTTTGTGTCAGGAAATGAGAGATTAGCTATAGGATTC 2626
Db 2429 CATCTTTAGATGAATGATTTGTTGTTAAGAAATGCGGTAAGGCTTTTAACATTC 2498
QY 2627 CATTTGTTGATTTGGAGGACCAACTCTCAAAACCCACACAGCAGTATAAATAG---CTC 2683
Db 2489 CTTTATTAATTTGGTGGCGGACTACTTCTTAAGCAATACAGCAGTAAATAATTGACCTC 2548
QY 2684 CGAGATACAGTGCACCTGTAATCCATGTCCTGGACGGTCCAAAGAGTGTGGTGTGTT 2743
Db 2549 AGTATCAAAACGATCGGTAATTTATGTTGCCGATGCTTCACTGCTGTGTGTAGCGA 2608
QY 2744 CCCAGCTGTAGTGAATAATCTAAAGGATGATCTTTGAGGAATCATGGAAGATATG 2803
Db 2609 CAACCTTCTTTGCAAGAAATGCGTGGAGCATTTATTTGAAGAGCATCGTCTGAATATG 2668
QY 2804 AAGATATTAGACGAGCACTATTAGTCTCTCAAGGAGGAGATCTTACCTTTAGTC 2863
Db 2669 CCAAAATTCGTGAGCTTTAGCCAAACAAACCAAAAGCAGCCAACTGACTTATAAG 2728
QY 2864 AAGCAGAAAAGTGGTTTCCAAATGGAATGGCTGTCTGAACTCTCACCCAGTGAAGCCCA 2923
Db 2729 AGTCGGTTGAAAATGTTTTAAATTTGATGAAGTACGTGCCAC-----CAAAACCA 2782
QY 2924 TGTTTATGGGACCCAGTCTTTGAAAGTATGATGACCTGCGAAGCTGTGACTACATG 2983
Db 2783 ATCTTTTGGGAACAACAAGTTTAAAGAAATATCCGCTGTCTACACTCGTGGATATTG 2842
QY 2984 ACTGAGGCTTCTTCTGATGCTGGCAGCTCCGGGGCAAGTACCCGATCGAGGCTCC 3043
Db 2843 ACTGAGCCCACTTCTTTATTTCTTGGAGTTTAACTGGCAA-----ATTCC 2887
QY 3044 CCAAGATATTAAACGACAAACAGTAGGTGAGAGGCCAGGAGCTCTACGATGATGCC 3103
Db 2888 CGAAAATTTTAGAGATGAGTGTGCGCGAAGCAGCAACTGACCTTTTACAAACAGGCAC 2947
QY 3104 ACAATATGCTGAACACACTGATTTAGTCAAAAGAACTCCGGGGCCGGGGTGTGGGT 3163
Db 2948 AAGCGATGTTGAAAGATATTATCGACAACTATCGTTTGTGATGCTCGTCTGTATTGTA 3007
QY 3164 TCTGCCAGCACAGATGATCCAGACGACATTCACCTGTACGACGAGCTGTGCCCC 3223
Db 3008 TGTTCCCTGCTCAGGCTACAGATGACGATCCGTCAGCGTATTTGATGAACCTGTCAA 3067
QY 3224 AGGCTGACAGGCCATAGCCACTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTG 3283
Db 3068 ATGTTACGCA-----TACTTTGAGCACTAGCCAGCAATCTGCAAGGTGACAG 3118
QY 3284 CCAGCACGGAGCCATPACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCC 3343
Db 3119 GCA-----AACCAAAATTTATCTTGGCAGATTACAT---TCGAGCTGATCGCGAGC 3169
QY 3344 GTGACTACTGGGCTGTTTGGCGTGTGCTGCTTGGGTGAGAGAGCTGAGCAAGGCT 3403
Db 3170 AAGACTACTTGGGCGGATTTCACTGTATCGATTTTGGTGCAGAAAGAACTGGCAATGA 3229
QY 3404 ATGAGGATGATGGTGAACACTACAGCAGCATCATGCTCAAGCGCTGGGGACCGGTGG 3463
Db 3230 ACAAGCCAAAGGTGATGACTACTCTGCAATTTTAGTCAGTCAATAGCTGACCGTTTG 3289
```

```
QY 3464 CAGAGCCTTTGCAAGAGCTCCATGAAAGAGTTTCGCGAGAACTGTGGCCCTACTGTG 3523
Db 3290 CCGAAGCTTTTGGCCGAACACTTACATGACGATATTTCGTAAGAGTTCTTGGGGCTATAAG 3349
QY 3524 GCAGTGACGAGCTGAGCTGCGACACCTGCGAAGTTTGGGTACAAAGGCAATCCGCCGG 3583
Db 3350 CTGATGAGCAGCTCAGCAATGAGAACTGATTAAGAGAAATATGTGGTATTTCGCCCTG 3409
QY 3584 CTCTGCTGCTACCCAGCCAGCCGACCAACCGAGAAAGCTCACCATGTGGAGACTCGCAG 3643
Db 3410 CACCAGGCTACCTGCTTGGCCAGAGCACTCTGAAAAAGCAGTGTGTTCTGACTGGTTAG 3469
QY 3644 ACATCGAGCAGCTCTACAGGCAATAGGTTAAAGAACTATAGCAATGCGACCTGCTTCAG 3703
Db 3470 GTTCTACCGAATAATCGGCACCAACTGACTGAGCACTTTGCAATGATGCCGCATCTT 3529
QY 3704 CAGTCTCAGGCTCTACTCTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTT 3763
Db 3530 CAGTAAGTGGTTCTATTATTCTCTCCTCAGAGTGAATACTTTAACTGGGTAAATCT 3589
QY 3764 CCAAGGATCAGGTTGAGGATTTGCAATTTGAGGAAGAAACATATCTGTGCTGAGTTGAGA 3823
Db 3590 CTCAAGACCAACTTGAAGATTATGCAAAACGTAAGGTTGGACACTGGATGAAGCGAAGC 3649
QY 3824 AATGGCTTGGACCACTTTTGG 3844
Db 3650 GTTGGTTAGTCCGAATTTAG 3670
```

## RESULT 10

ABD14844/c

ID ABD14844 standard; DNA; 3762 BP.

XX AC ABD14844;

XX XX 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polynucleotide #13448.

XX XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;

XX KW antibacterial.

XX XX Pseudomonas aeruginosa.

XX OS US651795-B1.

XX PN 22-APR-2003.

XX PD 18-FEB-1999; 99US-00252991.

XX PF 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX XX (GENO-) GENOME THERAPEUTICS CORP.

XX PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX PI WPI; 2003-615309/58.

XX DR P-PSDB; ABO81273.

XX DR Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of

XX PT pathological conditions resulting from bacterial infection.

XX XX Disclosure; SEQ ID NO 13448; 455pp; English.

XX PS The invention relates to Pseudomonas aeruginosa polypeptides and the

XX CC polynucleotides encoding them. The sequences are useful in diagnosis and

XX CC therapy of pathological conditions, as molecular targets for diagnostics,

XX CC prophylaxis and treatment of pathological conditions resulting from a

XX CC bacterial infection, for evaluating a compound, such as a polypeptide,

XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of

XX CC effective antibacterial targets, as targets for antibacterial drugs,

including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABD01397-ABD17967 represent P. aeruginosa polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

xx SQ Sequence 3762 BP; 548 A; 1172 C; 1294 G; 748 T; 0 U; 0 Other;

Query Match 23.0%; Score 903; DB 11; Length 3762;

Best Local Similarity 55.2%; Pred. No. 5.4e-246;

Matches 2074; Conservative 0; Mismatches 1580; Indels 103; Gaps 12;

Qy	113	CCCTGCGGGATGAGATCAATGCCATTCTGCAGAGAGAGGATATGCTGCGATGAGGGA	172
Db	3701	CCCGCTGCAAGCCTCCAGCACGCCCTCAGGGAACGTATCTGATCCTGATGCGGCA	3642
Qy	173	TGGGACCATGATCCAGCGGAGAGCTTAACGAGAACACTTCGGAGTTCAGGAATTA	232
Db	3641	TGGGACCATGATCCAGACTACAGCTGGAAGAGCGGACCTACCGGCGAGCGCTTG	3582
Qy	233	AAGATCATGCCAGCGCGCTGAAAGGCAACATGACATTTTAAGTATTAACCTCAGCTGATG	292
Db	3581	CGAGTGGCCGAGGAGCGTGAAGGCAACAGACCTCTTCTGCTGAGCGCGCCGAGCG	3522
Qy	293	TCATTACCAATCCATAAGGAATATCTTGCTGGCTGGGCGAGATATCATTTGAAACAAATA	352
Db	3521	TGATCCAGGCCATCGAGAAGCCTACCTCGACGCGCGCGCGACATCTCTCGAGACCAACA	3462
Qy	353	CTTTAGCAGCACTAGTATTCGCCAGCTGACTATGGCTTGACACTTGGCTTACCGGA	412
Db	3461	CTTTCAACGCCACCCAGGTGTCCAGGCGGCTACGCGATCAGTCTGCTGGCTACGAAC	3402
Qy	413	TGAACATGTCTGTCAGAGTGGCCAGAAAGCTGCCAGAGGTAACTCTCAGA---	469
Db	3401	TCALCTCGAAGGGCGCGCTTGGCCCGCAGGTGGCGAGCGAGAACCCCGGAGCCC	3342
Qy	470	CAGGAATTAAGAGGTTTGTGGCAGGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGT	529
Db	3341	CGGCAAGCCGCGCTTGTGCGCGGCGTGTGGCCCGACCGCGCACCTGCTCGATTT	3282
Qy	530	CCCATCTGTGGAAGGGCGGATTAAGGAACATCAATTTGATGAGCTTTGTAAGCAT	589
Db	3281	CCCCGGAGTGAACAAACCCGCGCTACCGCAACGTCACCTTCGACGAACCTGTTGGAGAACT	3222
Qy	590	ACCAAGACGAGCCAAAGAGCTTCTGGATGGCGGTTTGATATCTTACTCATTTGAAACTA	649
Db	3221	ACGTGAGGCGACCCGCGGCTTGATCGAAGCGCGCGCGACCTGATCTGTATCGAGACCA	3162
Qy	650	TTTTTGATCTGCCAATGCCAAGGCGAGCTTTGTTGCACTCCAAATCTTTTGGAGGAGA	709
Db	3161	TCTTCGACACCTCAACGCCAAGGCGGCGATCTTCGCGCTCCAGGCGGTGTTCCAGGAAC	3102
Qy	710	AATATGCTCCCGGCTATCTTTATTCAGGAGCATGTTGATTAAGTGGCGGACTC	769
Db	3101	TGGGCTGGAGTGCATGATCTCCGGAACCATCACCAGCGCTCCCGCGCACCC	3042
Qy	770	TTTTCGGGACAGACGAGGAGGATTTGTATCAGCGGTGCTCATGGAGAACCACTCTGCA	829
Db	3041	TGTCCGGGCGAGCACCGAGGCTTCTGGAATCTGGTGGGATGCCCGGCGATCTCGG	2982
Qy	830	TTGGATTAAATTTGCTTTGGGTGACGTGAGATGAGACCTTTTATTGAAATTAATTGAA	889
Db	2981	TAGGCTGAACTGCGCCCTCGCGCCAAAGGAATTCGGCGGTACATCGAGGAATGTGCA	2922
Qy	890	AATGTACAAACAGCTATGCTCTCTGTTATCCAAATGCAAGTCTTCCCAACACCTTTGGTG	949
Db	2921	CCAAAGCCGACACTATGTCTCGGCGCCACCCCAACCGCGGCTGCGCAACGCTTTCCGGC	2862
Qy	950	ACTATGATGAACCGCTTCTATGATGGCCCAAGCACCTTAAGGATTTTGTATGGATGGCT	1009

Db	2861	AATACGACGAATCGCGCGGAAATGCGCTGTGTGTCGAGGAATTCGCGCGCGCGCT	2802
Qy	1010	TGCTCAATATAGTTGGAGGATGCTGTGGTCAACACAGATCATATCAGGGAATTTCTG	1069
Db	2801	TCCTCAATATCGTGGCGGCTGTGGGCAACCCCGGCGACATCAGGCGGATCGCCA	2742
Qy	1070	AAGCTGTGAAATTTGAGCCTAGAGTTCCACTGCTGCACCTGCTTTTGAAGGACATATGT	1129
Db	2741	AGGAGTGGCAAGTACCCGCGCGGCGCATCCGGAGATTCCCGGCGCTGTC-----	2688
Qy	1130	TACTGTCTGTCTAGAGCCTTTCAGGATTCGACCGTACACAACTTTTGTAACTTTGGAG	1189
Db	2687	GCCTGTCGGGCTTGGAGCGCTTCAACATCGACCGAGCTGCTGTCTCAACGTCGCGC	2628
Qy	1190	AGCGCTATATGTTGAGGATCAAGGAGTTTCTAACTCATCATCGGAGGAACATATG	1249
Db	2627	AGCGCACCAATCACCGGTTTCGGCCAAAGTTTCGCCCGGCTGATCCGCGAGGAACATCG	2568
Qy	1250	AAGAAGCTTGTGTGTCCAAAGTGCAGGTGGAATGGGAGCCAGGTGTGATGTCA	1309
Db	2567	CGGAAGCTCTCGAGTTCGCCAGCAGCAGGTGGAAGCGCGGCCAGGTGATCGACATCA	2508
Qy	1310	ACATGGATGATGGCATGCTAGTTCAGTGCATGCAATGACACAGATTTTGCACCTTAATG	1369
Db	2507	ACATGGAGGAAGCATGCTGCACTCGAAGCGCGCATGCTCACTTCTCAACCTGATCG	2448
Qy	1370	CTTCGAGCAGACATCGCAAGGTACTTTTGTGATCGACTCTCCAAATTTTGTCTGTA	1429
Db	2447	CCTCGAGCCGACATCTCGCGTGCCTGATCATGATCGACTCTCTCAAGTGGGAAGTGA	2388
Qy	1430	TTGAAGCTGGTTTAAAGTGTGCCAAGGGAAGTGCATTTGTCAATAGCATTAGTCTGAAG	1489
Db	2387	TCGAGCGCGCTGAGTGTATCCAGGCGAAGGCACTGCTGACTCTGATCTCGATGAGG	2328
Qy	1490	AAGGAGAGCAGCTTCTTGGAGAGGCCAGGAAGATTAAAGTATGGAAGTCTGCTATGG	1549
Db	2327	AAGCGTCGAGGCTTCAAGCACCATGCCCGCTGTGCAAGCGCTACGGCGCGCGTGG	2268
Qy	1550	TGTCATGCTTTTGAAGAGGACAGGACAGCAACAGAAACAGACACAAATCAGAGTGT	1609
Db	2267	TGTCATGCGCTTCGACGAGCGCGCGCGCGCACCCAGGCGCGCAAGGAATCT	2208
Qy	1610	GCACCCGCGCTACCATCTGCTGTGAAATAAACTGGGCTTTAAATCCAAATGACATATTT	1669
Db	2207	GCAAGCGCTCTCAGCATCTGTCGAGAGTTCGCTTCCACACCGGAGACATCATCT	2148
Qy	1670	TTGACCTTAATATCTTAACCATTTGGGACTGGAATGGAGGAACACAACTTGTATGCCATTA	1729
Db	2147	TCGATCGCAATCTTCGCCCATGCCACCGCATCGAGGAACACAACTACCGCGTGC	2088
Qy	1730	ATTTTATCCATGCAACAAAGTCAATTAAGAAACATTTACTGGAGCCAGATAAGTGGAG	1789
Db	2087	ATTTTATCAACGCTTGGCGCTTACATCCGCGACCACTGCCCTACGCGCTCGGCGC	2028
Qy	1790	GTCTTTTCAACTGCTCTTCTTCGAGGAATGGAGCCATTCGAGAAGCAATTCGATG	1849
Db	2027	GGGTGTCAGAGTGTCTTCTGTTTCGCGCAACAACCCCGTGGCGGAGCGATCCACT	1968
Qy	1850	GGGTTCCTTTTACCATGCAATCAAGTCTGCGCATGGAATGAGATAGTGAATGCTGAA	1909
Db	1967	CGGTGTTCTCTACTACCGCATCCGCAACGCGCTGACCATGGGCATCGTCAACGCGCGC	1908
Qy	1910	ACCTCCCTGTGTATGATGATATCCATAGGAATCTTCTGAGCTCTGTGGAATCTCATCT	1969
Db	1907	AGCTGGAATCTACGACGAGATTCGGAAGCGCTGCGCGACCGGTTCAGGACGTTGGTGC	1848
Qy	1970	GGAATAAGACCTTGAGGCCACTGAGAAGCTTTAGTTATGCCAGACTCAA---GGCA	2026
Db	1847	TCACCGCACCGCGGCGCACCGAGGCTTGTGGCGATCGCGACACTACAAGGCGG	1788
Qy	2027	CAGGAGGAAGAAAGTCAATTCAGATGATGAGTGGAGAAATGGCCCTGTGGAAGACGCC	2086



XX (GENO-) GENOME THERAPEUTICS CORP.  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI; 2003-615309/58.  
XX P-PSDB; ABO81518.  
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX Disclosure; SEQ ID NO 13693; 455pp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-*P. aeruginosa* drugs, as templates for recombinant  
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused  
XX infection, and in detection of *P. aeruginosa* sequences or other sequences  
XX of *Pseudomonas* species using biochip technology. Sequences ABD01397-  
XX ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 3795 BP; 751 A; 1307 C; 1177 G; 560 T; 0 U; 0 Other;  
XX  
XX Query Match 23.0%; Score 903; DB 11; Length 3795;  
XX Best Local Similarity 55.2%; Pred. No. 5.5e-246;  
XX Matches 2074; Conservative 0; Mismatches 1580; Indels 103; Gaps 12;  
XX  
QY 113 CCCTCGGGATGAGTCAATGCCATCTGCGAGAGAGATTATGGTGTGGATGGAGGA 172  
DB 119 CCCGCTGCAAGCCCTCCAGACGCCCTTCAGGGAACGTATCCTGATCTCGATGGCGCA 178  
QY 173 TGGGACCATGATCCAGCGGAGAGCTTAAACGAGAACACTTCCGAGGTGAGAAATTTA 232  
DB 179 TGGGACCATGATCCAGAGCTACAAGCTGGAAGGCCGACTACCGCGGCGAGCGCTTCG 238  
QY 233 AAGATCATGCCAGGCGCTGAAAGGCAACAATGACATTTTAAGTATTAAGTCACTGATG 292  
DB 239 CCGACTGGCCGAGCGACGTGAAAGGCAACAACGACCTCTTGTCTGAGCGCGCCGACG 298  
QY 293 TCATTTACCAATCCATAAGGAATCTTGTGGCTGGGCGAGATATCATTTGAAACAATA 352  
DB 299 TGATCCAGGCCATCGAAGAGCCCTACTTCGACGCGCGCGCGGACATCTCGAGACCAACA 358  
QY 353 CTTTATGACGACATAGTATGCCCAAGCTGACTATGGCTTTGAAACACTTGGCCCTACCGGA 412  
DB 359 CTTTCAAGCCACCCAGGTGTCCAGGCCGCTACGAGTACGATCGCTGCGCTTACGAC 418  
QY 413 TGAACATGTGCTTGCAGAGGTGGCCAGAAAGCTGCCAGAGGTAACTCTTCAGA --- 469  
DB 419 TCAACCTCGAAGGGGCGCGCTTGGCCCGCAGGTGGCGAGCGGAAGACCCCGCAGACCC 478  
QY 470 CAGGAATTAAGAGTTTGTGACAGGGCTCTGGGTTCGACTAATAAGACATCTCTGTGT 529  
DB 479 CGGAACAGCCGCGCTTGTGCGCGCGCTGTCTGGCGCGCGACCGCGCACCTTGCTGATTT 538  
QY 530 CCCCATCTGTGGAAGGCGGATATAGGAACATCACATTTTGTGAGCTTGTGTAAGCAT 589  
DB 539 CCCCAGAGCTGAACAAACCCCGGCTACCGCAAGCTCACCTTCGACGAATCTGTGGAGACT 598  
QY 590 ACCAAGAGCAGGCCAAAGGACTTCTGGATGCGGGGTTGATATCTTACTTAAAGACTA 649  
DB 599 ACGTGAGGCGACCGCGCGCTGTGTAAGAGCGCGCGGACCTGATCTGTATCGAGACCA 658

QY 650 TTTTGTATCTGCCAATGCGCAAGCGACCTTGTGTCACCTCCAAAATCTTTTGGAGGA 709  
DB 659 TCTTCGACACCTCAAGCCCAAGCGCGGATCTTCGCGCTCCAGGGGCTGTTCAGGAA 718  
QY 710 AATATGCTCCCGGCTATCTTTATTTTCAAGGACGATCGTTGTGATAAAGTGGGCGGACTC 769  
DB 719 TCGGCGTGGAGCTGCGGATCATGATCTCCGGAACCATCACCGACGCTCCGCGCGCACCC 778  
QY 770 TTTCCGACAGACGAGGAGGGATTTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCA 829  
DB 779 TGTGCGGCCAGACACCGAGGCTTCTGGAACCTCGGTGCGGCATGCCCGCGCATCTCGG 838  
QY 830 TTGGATTAAATTTGCTTTTGGGTGCGAGCTGAGACCTTTTATTGAAATTAATGGAA 889  
DB 839 TAGGCTGAATGCGCCTCGCGCCCAAGAAATTGCGCGGTATCATCGAGAACTGTGCA 898  
QY 890 AATGTACAACAGCCTATGCTCTGTTATCCCAATGAGGTCTTCCAAACCTTTGGTG 949  
DB 899 CCAAGGCCGACACTCATGCTCTCGGCCCAACCCCAAGCGCGCTGCCGAGCCCTTCGGCG 958  
QY 950 ACTATGATGAAACGCTTCTATGATGGCCCAAGCACCTTAAAGGATTTTGTCTATGATGCT 1009  
DB 959 AATACGACGAATCGCGCGCGGAATGCGCTGTGTGAGGAATTCGCGCGCGCGCT 1018  
QY 1010 TGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACACAGATCATATCAGGGAATTTGCTG 1069  
DB 1019 TCCTCAATATCTCGCGGCTGTTCGGCACCAACCGCGCGCATCGAGCGGATCGCCA 1078  
QY 1070 AAGCTGTGAAAAATTTGAAAGCTGAGTTCCTGCTGCACTGCTTTTGAAGGACATATGT 1129  
DB 1079 AGGCACTGGCCAAAGTACCCTCGCGGCCCATCCCGAGATTCCCGCGCGCTGTC ----- 1132  
QY 1130 TACTGTCTGTCTAGAGCCCTTCAGGATTTGGACCGTACCAACTTTGTTAACTGGAG 1189  
DB 1133 GCCTGTCCGCTCGAGCGCTTCAACATCGACCGAGCTCGCTGTTCGTCAACGTGCGCG 1192  
QY 1190 AGCGCTTAATTTGCGAGGATCAAGGAATTTGCTAACTCATCATCGGAGGAACTATG 1249  
DB 1193 AGCGACCAACATCACCGTTTCGCCCAAGTTTCGCCGCTGTATCCGCGAGGAACTACG 1252  
QY 1250 AAGAAGCTTTGTGTTTGCACAAAGTGCAGGTGGAAATGGGAGCCCAAGGTGTGATGTC 1309  
DB 1253 CGGAAGCTCTCGAGTCCGCCAGCAGCAGGTGGAAGCGCGCGCGCTGATGCACATCA 1312  
QY 1310 ACATGATGATGCGATGCTAGATGGTCCAAGTGCATGACGAGTTTGGAACTTATTTG 1369  
DB 1313 ACATGGAAGGAGGATGCTGGAATCGAAGCGGCCATGGTCACTTCTCAACCTGATCG 1372  
QY 1370 CTTCCGAGCCAGACATCGCAAAAGGTACTTTTGTGATCGACTCTCTCAAATTTTCTGTGA 1429  
DB 1373 CTTCCGAGCCGACATCTCGCGGTGCGGATCATGATCGACTCTCTCAAGTGGGAAGTGA 1432  
QY 1430 TTGAAGCTGGGTTAAAGTGTGCCAAGGGAAGTGCATTTGTCAATGACATTAAGTCTGAAG 1489  
DB 1433 TCGAGGCGCGCTGAGTGCATCCAGGCGCAAGGSCATCGTCACTCGATTCGATGAAG 1492  
QY 1490 AAGGAGAGGACGACTTCTTTGGAGAGGCGCAGGAAGATTAAAGATATGAGAGCTCTATGG 1549  
DB 1493 AAGCGCTCGAGGCGCTTCAAGCAACCATCCGCGCTGTGCAAGCGCTACGCGCGCGGTGG 1552  
QY 1550 TGGTCATGGCTTTTGTATGAAGAGGACAGGCAACAGAAAACAGACACAAAATCAGAGTGT 1609  
DB 1553 TGGTATGGCTTCGACGAGGACGCGCGCGCGGACCCGAGCGCGCGGAGGAAATCT 1612  
QY 1610 GCACCCCGGCGCTACCTCTGTGTGAAAAAACTGGGCTTTTAATCCAAATGACATTAATTT 1669  
DB 1613 GCAAGCGCTCCTACGACATCTTGTTCGACGAGTTCGGCTTCCCAACCGGAGACATCTCT 1672  
QY 1670 TTGACCTTAATATCTTAACCATTTGGGACTGGAATGGAGGAGAACACAACTTTGATGCCATTA 1729  
DB 1673 TCGATGCGAAACATCTTCGCCATCGCCACCGGCATCGAGGAGAACACAACTACCGGTCG 1732  
QY 1730 ATTTTATCCATGCAACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATTAAGTGAG 1789

1733 ATTTCAATCAACGGCTGCGCTACATCCGCGACCACTGCGCTACGCGCTGAGCTCGGCG 1792  
1790 GTCTTTCCAACTTGCTCTCTCTCCGAGGAATGGAAGCAATTCGAGAAAGCAATGATG 1849  
1793 GGGTGTCCAAAGTGTCTCTCTGTTCCGCGCAACAACCCGGTGGCGAGCGATCCACT 1852  
1850 GGGTTTCTTTACATGCAATCAAGTCTCGGATGGAATGAGATAGTGAATGCTGGAA 1909  
1853 CGGTGTCTCTCTACTACGCAATCCGCAACGCGCTGACCATGGGCATCGTCAACGCGGCC 1912  
1910 ACTCTCCCTGTGATGATATCAATGAAGCAATTCGCAAGCTCTGTGAAGATCTCATCT 1969  
1913 AGCTGGAAATCTACGACGAGATTCGAAAGCGCTGCGCGACCGGCTCGAGAGCTGGTGC 1972  
1970 GGAATAAAGCACTTGAGGCACTGAGAAGCTCTTACGTTATGCGCCAGACTCAA----GGCA 2026  
1973 TCAACCGCACGCCGAGGCCACCGAGGCCCTGTGCGGATCGCGGACGACTACAGGGCG 2032  
2027 CAGAGGGAAGAAAGTCAITTCAGACTGATGAGTGGAAATGCGCCCTGTGGAAGAGGCC 2086  
2033 CGCGCGCGGTCAAGGAGCGCGAGGACGAGGAATGGCGCAGCTACAGTGTGGAAGCGCC 2092  
2087 TTGAGTATGCCCTGTGGAAGGCAATTCGAAACATATATTGAGGATCTGAGGAGCCA 2146  
2093 TCAGAGCATGGCGTGGTCAAGGGGCATCCACCTGGATCGTGGAGCACCGCGA----- 2147  
2147 GGTAAACCAAAAAAATATCCCGGACCTCTCAATATAATTGAAGGACCCCTGATGAATG 2206  
2148 -----ATCGCGCAGCAGTGTGCGGCTCCATCGAGGTCTCGAAGGTCCGCTGATGCC 2203  
2207 GAATGAAATTTGTTGGTGTATTTTTGGAGCTGGAAATATGTTTCTACCTCAGGTTATAA 2266  
2204 GGATGAACGTGTCGCGGACCTGTTTCGCGCGCGGCAAGATGTTCTCCCGCAGGTGTCA 2263  
2267 AGTCAGCCCGGTTATGAAGAGGCTGTTGGCAGCTTTATCCCTTTCATGCAAAAGAAA 2326  
2264 AGTCGCGCGAGTGATGAACACAGGCGGTGGCCACCTGATTCCTCTTCATCGAGCGGGA 2323  
2327 GAGAAGAACACAGAGTGTCTAACGGCACATAGAGAAGAGGACCTTACACAGGGCACCA 2386  
2324 AAGCGGATA-----AGCCGGAAGCCAAAGGGCAAGA 2353  
2387 TCCTGCTGGCCACTGTTAAAGGCGACGTGCACGACATAGGCAAGAAATAGTGGAGTAG 2446  
2354 TCCTGATGGCCACGCTGAAGGCGACGTGCAGGACATCGGCAAGAAATCGTGGCGTGG 2413  
2447 TCCTTGGCTGCAATATTTCCGAGTTATGATTTAGGATCATGACTCCATGTGATAGA 2506  
2414 TGCTCGGCTGCAACGGCTATGACGTGGTTCGACCTCGGCGTGTGTTGTCGCGGAGAGA 2473  
2507 TACTGAAGCTGCTTTCACCAAGACGATATAATTGGCTGTGAGGACTCATCACTC 2566  
2474 TCCTGACAGCGCCATCGCGAGAAATTCGACATCATCGGCTGTCTGGCTGATCAGC 2533  
2567 CTTCCTCGGATGAATGATTTTGTGCAAGGAAATGGAGAGATTAGCTATAGGATTC 2626  
2534 CGTGGCTGGAAGAGTGTCCAGCTGCGCAAGGAATTCAGCGCGAGAAATTCAGTTGC 2593  
2627 CATTTGTTGATGGAGGAGCAACCACTTCAAAACCCCAACAGCAGTTTAAATAGTCCGA 2686  
2594 CGCTGATGATCGGCGGCGCACTACCTCGAAGGCGCATACCGGCTGAAGATCGATCCGC 2653  
2687 GATACAG-----TGCACTGTGTAATTCATGCTTGGAGCGGTCCAGAGTGTGGTGTGTT 2743  
2654 AGTACAGCAACGACGCGGTGTCTTACGTTCACCGACGCTCGCGCGCGGTAGGCGGCA 2713  
2744 CCCAGCTGTTAGATGAATAATCAAGGATGAATGATTTGAGGAATCATGCAAGAAATG 2803  
2714 CCAGCCTGCTGTCGAAGAGCTGAAGCGCGGACTACGTGGGCGCCGCCACCGCGCGACTACG 2773  
2804 AAGATAT-----TAGACAGGACCAATTATGATCTCTCAAGGAGAGGAGATCTTACCCCTT 2858

RESULT 12  
ACA44722

2774 CGGTGCTCGCGNAACGACGCGCAACCGCAGCGCCCGCACCGAGCGGCTGAGCTACGAAC 2833  
2859 AAGTCAAGCCAGAAAAAGTGGTTTCAAAATGGATGGCTGTCTGAACCTCACCAAGTGAA 2918  
2834 AGCGGATCGCAACAAGCGCGCTTCGACTGGCGCGCTACCAAGGCGCGAC----- 2885  
2919 GCCCAGTTTATTTGGGACCCAGGTCTTTGAAGACTATGACCTGCAGAGCTGTGGACTA 2978  
2886 GCCTTCTTTCACCGCGCTCAGGGTGTCTCGACGAGATCGACCTCGCGGTGTCTCCCGAGTA 2945  
2979 CATTTGACTGAAGCCTTTCTTTCATGCTCTGGCAGCTCCGCGGCAAGTACCCGAATCGAG 3038  
2946 CATCGACTGGAACCGCTTCTTCATTTCTTGGGACCTTGGCGCGCAAGTACCCG----- 2997  
3039 CTTTCCCAAGATATTTAAACGAAAAACAGTAGTAGTGAGAGCGCCAGGAAAGTCTACGATGA 3098  
2998 -----CGCATCTCACCGACGAAGTGTGCGGAGGCGCCACCTCGTGTGTTCAACGA 3050  
3099 TGCCCACAATATGCTGNAACACTGATTAAGTCAAAAGAAACTCGGCGCCCGGGTGTGGT 3158  
3051 CGCCCAAGGCGATCTGAAGAAGCTGATCGACGAGAGCTGATCAAGGCGCCGCGGTGTT 3110  
3159 TGGGTTCTTGCCGACGACAGAGTATCAAGACGACATTCACCTGTATCGCAGAGGCTGCTGT 3218  
3111 CGGCTTCTTGCCCGGCCAACAGAGTTCAGACACGACGACCTCGAGAGTCTACGCGCCGATGG 3170  
3219 GCCCAGGCTGACAGAGCCCATAGCCACTTTTCTATGGGTTAAGGCAACAGGCTGAGAAAGGA 3278  
3171 C-----GAGACCTTCGCCACCTGCACCTCGCGCAGCAGACGATCAAGCC 3218  
3279 CTCTGCAGCAGCGGACGATACCTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGG 3338  
3219 GGAAGGCA-----AGCCGAACCTGCTGCTGGCGATTTCTGTCGCGCGGAAAGAACGG 3272  
3339 CATCTGTTGACTACTCGGCGCTGTTG---CGTTGCTGCTTTGGGGTAGAAGCTGAG 3395  
3273 CGTGGCGGACTACATCGCGGCTTTCATCACACCGCGGATCGGCGCCGAGGAGTGGC 3332  
3396 CAAAGGCTTATGAGGATGATGTTGACGACTACAGCAGCATCATGTTCAAGCGCTGGGGGA 3455  
3333 CAAAGGCTACGAAGCCAAAGGCGACGACTACACAGCATCATGTTCAAGCGCTCGCGGA 3392  
3456 CCGGCTGGCAGAGCGCTTTGCAAGAGCTTCATGAAGAGTTTCGCGCAGAACTGTGGGC 3515  
3393 CCGCTTCGCGAAAGCTTCGCGGAGTGGCTGCAAGCGCGGTGCGCAAGGAGTACTGGGG 3452  
3516 CTACTGTGGCAGTGAAGCTGCAAGCTCGCAGACCTCGAAGGTTTCGGGTACAGGCGAT 3575  
3453 CTACGCGCGCAGCAGACCTTCGACCAACGAGGCTTTGATCAAGGAGCAATACGTCGGCAT 3512  
3576 CCGCCCGGCTCTCTGGCTTACCCAGCAGCCGACCAACAGGAGAGCTCACCATGTGGAG 3635  
3513 CCGCCCGGACCGGGCTACCCGCGCTTCGCGGACCATACCGAAGAGGACCTTGTTCGA 3572  
3636 ACTTCGACGACATCGACGAGTCTACAGGCAATTAGGTTAAACAGAAATCATTAAGCAATGGCACC 3695  
3573 ACTGCTGATTCGCGAGCGGCTGTCGCGGCTGACGCTGACCGACGACTACGCGATGTTCCC 3632  
3696 TGCTTCAGGATCTCAGGCGCTTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGG 3755  
3633 CGCGCGGCGGTGAGCGGTTGGTATTTTCGCCACCCCGCAGGCGCAGTACTTCGCGGTGG 3692  
3756 GAAGATTTCCAAAGGATCAGGTTGAGGATTTATGATTGAGGAAGAAACATATCTGTGGCTGA 3815  
3693 CAAGATGCAACAGGACGAGGTGGAAACGCTACAGCAGCGCAAGGCCAGGAAGCCAGCT 3752  
3816 GGTGAGAAATGGCTTGAGCCCATTTTGGGATATGAT 3852  
3753 CAGCGAGCGCTGGCTGGCGCGGAACTTGGCTAGAT 3789







Db 1048 GATAATTCGCTGTCGTCATGTCGGGGAACGAACGAATGTCACAGGTTCCGCTTAATTT 1107  
Qy 1222 GCTAAATCATATGCGAGGAAACATATGAAGAGCCCTTGTGTGTCCAAAGTGCAGGTG 1281  
Db 1108 AAACGTTTAAATAAAGAGGTAAATACCAAGAGCCCTTGATGTTGCTGTCGCAAGTT 1167  
Qy 1282 GAATGGAGCCAGGTTGGATGTCACATGATGATGCGATGTCAGTGTCCAGT 1341  
Db 1168 GAAACGGTGCTCAATCATCATGATCAATGATGAAGGCATGTTAGATGCGATAGAA 1227  
Qy 1342 GCAATGACCAATTTGCAACTTAATGCTTCCAGCCAGACATCGCAAGAGTACCTTTG 1401  
Db 1228 GCAATGACAGTCTTCTTAATCTTATTTGCGGGTGAACCCGATATGCAAGTCCGAT 1287  
Qy 1402 TGCATGACCTCTCCAAATTTTGTGATGAAGCTGGGTTAAAGTGTCTGCCAAGGGAAG 1461  
Db 1288 ATGATTGACTCGTCAAAATGSCAAAGTTATTGAGGAAGGATTAATGATTCAGGTA 1347  
Qy 1462 TGCATTTGATCAATGATGATGTCGAAGAGGAGAGACGACTTCTTGGAGAGGCCAGG 1521  
Db 1348 GGAATGCTCAATTCGATGAAAGAGGGGGAAGCCCTTTTATTTGAGCATGCAAAA 1407  
Qy 1522 AAGATTAAAGATGAGAGCTGCTATGCTGCTCATGCTTGTGATGAAGAGGACAGCA 1581  
Db 1408 TTAGTTGATAGTATGCTGCAAGTTGTTGATGGCAATTTGATGAAGTATGGAAGCAAGCA 1467  
Qy 1582 ACAGAAACAGACACAAAATCAGAGTGTGACCCGGGCTACCATCTGCTTGTGAAAAA 1641  
Db 1468 GATACTCGAGAGCGAAAATAGAAATTTGTCGGGCTGGGTATCATCTATTAACTGA 1527  
Qy 1642 CTGGGCTTTAATCAAAATGACATTAATTTTGAACCTTAATATCTTAACATTTGGAATGGA 1701  
Db 1528 GCCGGAATTCACCAAGAGATTAATTTTGTGATCCTAATATTTTGTGTGAGCAACGGT 1587  
Qy 1702 ATGAGGACACAACTTGTATGCAATTAATTTTATCCATGCAACAAAGTCATTAAGAA 1761  
Db 1588 ATTGCTGAACATAATAATTTATGCGGTTGATTTATAGAAGTTTGTGCGGATTAATAATCT 1647  
Qy 1762 ACATTACTGGAGCGCAATAAGTGGAGGTCCTTTCACAACTTGTCTTCTCCGAGGA 1821  
Db 1648 CAATGCTCTTATGCTTAAATTTCTGGCGGTGTTTCTAATGTGTCAATCTCTTTTCGGGGC 1707  
Qy 1822 ATGAAGCCATTCGAGAGCAATGATGCGGTTTTCTTTACCAATCAATCAAGTCTGGC 1881  
Db 1708 AATGATCCGTTCCGGAAGCGATCACTCCGTTTTCTCTATTACGCTGTAAAAAATGTT 1767  
Qy 1882 ATGACATGGAGATGATGATGCTGGAACCTCCCTGTGTATGATGATATCCATAAGAA 1941  
Db 1768 ATGGATATGGGATGTCATGCTGGGCAACTCGCTATTTATGACTCATATACCGGATGAA 1827  
Qy 1942 CTTCTGCACTCTGTGAAGATCTCATCTGGAATAAAGACCTGAGGCCACTGAGAAGCTC 2001  
Db 1828 CTAGCTAATGCGGTAGAGGATGTTATTTAATCGCCATGCAAGAAAGTACTGATTAATTA 1887  
Qy 2002 TTACGTTATGCCAGACTCAAGGC-----ACAGGAGGGAAGAAAGTCATTCAGACTGAT 2055  
Db 1888 CTGGCACTCGCAGAGCGTTATCGTGGTACTATAAAGTGAAGGAACGACACTCTACGCTTTGCG 1947  
Qy 2056 GAGTGGAGAAATGCCCTCTGCAAGAGCGCTTGAATGATGCTTGTGAAGGGCAATTTGAA 2115  
Db 1948 CAATGGCGCAATGGGAGGTGCAAAAGCGCTAGAGTACGCGTTAGTGAAGAGGATCAC 2007  
Qy 2116 AAACATATTTAGAGTACTGAGGAAGCCAGGTTAAACCAAAAAAATAATCCCGACCT 2175  
Db 2008 GAATTTATCATGAAGATACCAAG-----CGTGGCCGCAACAAAGCTTCAGTCCT 2058  
Qy 2176 CTCAATATTAATGAAGGACCCCTCATGAATGGAATGAAATTTGTTGGTGTATCTTTTGA 2235  
Db 2059 ATCGAAGTATGAGAGGCGCAATTAATGAATGGCATGAACACTGTGGCGGATCTGTTGGT 2118  
Qy 2236 GCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTATGAAGAGGCTGTT 2295

Db 2119 GAAGGTAAATGTTTACCCCAAGTAGTAATAATCGCACGGGTGATGAACAGGCCGTA 2178  
Qy 2296 GGCACACCTTATCCCTTTTCATGAGAAAAGAAAGAGAAACCAAGAGTGTCTTAACGGCACA 2355  
Db 2179 GCTTACCTTGAACCTTATATTTCAAGCGACAAAAAAGAGCGCAG----- 2222  
Qy 2356 GTAGAAGAGAGGACCTTACAGGCGACCATCGTCTGCGCCACTGTTAAAGCGACGTG 2415  
Db 2223 -----CTCGGAGGAAAAAGTGTACTGCGCAGACAGTAAAGGGAGATGTT 2265  
Qy 2416 CACGACATAGCGAAGAACATAGTTGAGTAGTCTTGGCTGCAATTAATTTCCGAGTTATT 2475  
Db 2266 CACGACATTTGCGCAAAATATTTGTTGAGTGTGTGTCAGTGTAACTATGAATATATC 2325  
Qy 2476 GATTAGAGTCAKAGACTCCATGTGATAAGATACTGAAAGCTGTCTTGAACCAAAAGCA 2535  
Db 2326 GATCTGGAGTCAATGTGTACCTTGGATATAAATCTTGCMAACAGCATTTGATGAACATGTT 2385  
Qy 2536 GATATAATTGGCTGTGACGACTCATCACTCTCTTCCCTGCGATGAATGTTTTGTGCGC 2595  
Db 2386 GATATTATTGGATTGTCTGGATTGATCACTCCCATCACTGGATGAGATGGTGAATGTTGCT 2445  
Qy 2596 AAGCAATGAGAGATTAGCTATAAGGATTCCATTTGTTGATTGGAGGAGCAACCACTTCA 2655  
Db 2446 AAAGAGATGGAAGACGCGGTTTTCTCTTCTTTAATGATTGGTGGCGACAACATCG 2505  
Qy 2656 AAAACCCACACAGCAGTTAAAATAGCTCCGAGATACAGTGCACCTGTAAATCCATGCTCTG 2715  
Db 2506 AAAGCGCACAGCGAGTAAATAATTGAGCCTAATTACAGCCACCAACCGTTTATGTGCG 2565  
Qy 2716 GACGCTCCAAGAGTGTGTGTGTTCCAGCTGTTCCAGCTGTTAGATGAATACTTAAGGATGAA 2775  
Db 2566 AATGCGCTCAAGAGCGGTGCGGCTGCTGCTATTAATTCGGCCACACAAAAGGCTGAT 2625  
Qy 2776 TACTTTGAGGAATCATGGAAGAAATATGAAGATATTAGACAGACCAATTAATGAGTCTCTC 2835  
Db 2626 TTTGTCGTHAAACACGCCGTGAATATGAAGTGGTACGTGACGATGTCGAGAAAAAAA 2685  
Qy 2836 AAGGAGAGGAGATVACTTACCTTTAAGTCAAGCCAGAAAAAGTGGTTTTCCAAATGGATTGG 2895  
Db 2686 CCGGTACGCTCCCGTTTCTTTAGCGAAAGCTCGTGCCAAATGCAATTGCGAGTTGACTGG 2745  
Qy 2896 CTGCTGAACCTCACCAAGTGAAGCCCACTTTTATTTGGGACCCAGGCTTTTGAAGACTAT 2955  
Db 2746 CAACACTATACGCCACCTAAGCCGAACCAATTAGGGGTACAACAAAGTTACAGCCAAATAT 2805  
Qy 2956 GACCTGACAGAGCTGTGTGACTACATGCTGGAAGCCTTTCTTTGATGCTGCGACGCTC 3015  
Db 2806 GAA-----ACTTTACGTGAGTATATCGACTGGACGCTTTCTTTATGACTTGTGCTATTG 2859  
Qy 3016 CCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAACGACAAAAACAGTAGGTGGA 3075  
Db 2860 GCGGGGAATATCCAGT-----ATATTAGAAGTGAAGTTGTTCGGAGAA 2904  
Qy 3076 GAGGCGAGGAAGTCTACGATGATGCCCAATATGCTGAACACACTGATTAGTCAAAAG 3135  
Db 2905 GAGTAAGACGCGTTTTTGACATGCGAATGCCATGTTGGATAAACTCAGCCGTGAAAAA 2964  
Qy 3136 AAACCTCGGGCCCGGGTGTGGTTCTGCGCCACAGAGTATCCAAAGACGACATT 3195  
Db 2965 TTATGACACCGAAAGGATATCGTTGGTTTATTCCTGCTAATCGCCTTGGTGAATGACATT 3024  
Qy 3196 CACTGTACGACAGGCTGTGTGCCCGGCTGACAGGCTGACAGCCATAGCCACTTTCTATGGG 3255  
Db 3025 ATTATTACCAAGTGAACGCGACAGCATG-----AGTTATTACGTTGTTGCCAT 3075  
Qy 3256 TTAAGGCAACAGGCTGAGAGGACTGTGCCAGCAGCGAGCCACTACTCTCCTCTCAGAC 3315  
Db 3076 TTAGCGCAACAAACAGAAAAAAGA-----GTACCCCTAACATTCTCTTAGCAGAC 3126  
Qy 3316 TTCAATGCTCCCTTGCATTTCTGCGATCCGTCGACTACTGCGGCTGTTTGGCGTTGCTGCG 3375  
Db 3127 TTTATTGCCCCCTGTGGATAGTGTCTTGCTGACTATTTTGGTGTCTTTTGGCAGTTACCGGC 3186

```
Qy 3376 TTGCGGGTAGAAGA---GCTGAGCAAGCGCTATGAGGATGATGCTGAGCACTACAGCAGC 3432
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3187 GGTCTAGAGAGATGACCTCGCTAATGGGTATGATTAATGCTCATGATTAACAATAA 3246
Qy 3433 ATCATGTCFAAGCGCTGGGGGACCGGCTGCAGAGCGCTTTGAGAGAGCTCCATGAA 3492
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3247 ATTATGGTAAAGCGCTATCCGATGCTTTCGAGAGCGCTTTTCCGAATATTTGCATCAA 3306
Qy 3493 AGAGTTCGCCAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGAGCTGCGACCTG 3552
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3307 CAAAGTGGAACCAAGATTTGGGGTTACAGTCTCTGACGAAGCGCTCTCTAATGACGAGTTG 3366
Qy 3553 CGAAGGTTGGGTCACAGGGCATCCGCCGCTCTCGCTACCCAGCCAGCCGACCCAC 3612
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3367 ATTAGAGAAATATCAAGAACAGCCTGCACCGGGATATCCGGTTCGCCAGAGCAC 3426
Qy 3613 ACCGAGAGCTCACCATGTGGAGACTCGACACATCGAGCAGTCTACAGGCAATTAGGTTA 3672
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3427 ACAGAAAAGCAAAAATTTGGCAACTGTAAATGTAGAAAATCGCATTTGGTATGAAGCTT 3486
Qy 3673 ACAGATCATTAGCAATGGCACTGCTTCAGCAGTCTCAGGCTCTACTTCTCCAATTG 3732
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3487 ACTGATGCTTACGCGATGTGGCGGGGCTTCGGTATCGGGTTGGTATTTTAGTCATCCT 3546
Qy 3733 AAGTCCAAATATTTGCTGTGGGAAGATTTCCAAAGGATCAGGTTGAGGATTTATGCATTG 3792
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3547 GAAAGTAATATTTTGGCTGCTCAANTTCAAAAAGATCAGTAGAAGATTACGCCAA 3606
Qy 3793 AGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTTTGGGATATGA 3851
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 3607 CGTCTGTATGAGTGTGAGTGAAGTTGAGGCTTGGTTAGCCCTTAATTTAGGTTATGA 3665
```

RESULT 13  
ACA42279

ID ACA42279 standard; DNA; 3705 BP.

XX ACA42279;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #23936.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.

XX Pseudomonas aeruginosa.

XX WO20027183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342993P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX MPI; 2003-029926/02.

XX P-P8DB; ABU38409.

XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 30149; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pst\_sequences

SQ Sequence 3705 BP; 737 A; 1269 C; 1161 G; 538 T; 0 U; 0 Other;

Query Match

Best Local Similarity 23.0%; Score 901.4; DB 8; Length 3705;

Matches 2073; Conservative 0; Mismatches 1581; Indels 103; Gaps 12;

```
Qy 113 CCCTGCGGATGAGATCAATGCCATCTCGAGAGAGAGATATGCTGCTGATGAGGGA 172
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 173 TGGGACCATGATCCAGCGGGGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTA 232
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 89 TGGGCACCATGATCCAGAGCTACAAAGCTGGAAGAGGCCGACTACCGGGGAGCGCTTCG 148
Qy 233 AAGATCATGCGCAGCGCTGAAAGGCAACAATGACATTTTAAAGTATTAATCAGCCTCATG 292
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 149 CCGACTGCGCGAGCGAGCTGAAAGGCAACAACGACCTCTTCTGCTGAGCGCCGCGAGC 208
Qy 293 TCATTTACCAATCCATAAGGAATACCTTGTGCTGGGGGAGATATATGTAAGAACAAATA 352
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 209 TGATCCAGGCCATCGAGAAGCGCTACCTCGAGCGCGCGCGAGCATCTCTGAGACCAACA 268
Qy 353 CTTTTAGCAGCATAGTATTTGCCCAAGCTGACTATGCTTGAACACTTGGCCTTACCGGA 412
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 269 CTTTCAAAGCCACCCAGGTTGTCCAGGCGGACTACCGGCATGACGCTCGCTGCGCTACGAAC 328
Qy 413 TGAACATGTGCTCTGAGGAGTGGCCAGAAAAGCTGCGGAGAGGTAACTCTCCAGA--- 469
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 329 TCAACGTGGAAGGGCGCGCTTGCCCGCCAGGTGGGGGAGCGGAGAACGCCGAGACCC 388
Qy 470 CAGGAATTAAGAGGTTTGTGGCAGGGGCTCTGGGTCGAGCTTAATAAGACACTCTCTGTGT 529
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 389 CGGACAAAGCGGCTTTCGTGCGCGGCTGCTCGGCGCGGACCGAGCGGCTGCTCGATTT 448
Qy 530 CCCCATCTGTGGAAGCGCGGATTAAGGAACATCAATTTGATGAGCTTGTGAAGCAT 589
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dy 449 CCCCGGAGCTGGAACAACCCCGGCTACCGCAACGCTCACCTTCGACGAACTGGTGAGAACT 508
Qy 590 ACCAAGAGCAGCCCAAGGACTTCTGATGCGGGGTTGATATCTTACTCATTGAAACTA 649
```

[illegible]

Db	1583	TCGATCGGAACATCTTCGCGCATCGCCACCGGCATCGAGGAAACAACAACACTCAGCGGTGCG	1644
Qy	1730	ATTTTATCCATCAACAATAAGTCATTAAAGAAAATTAATCTCGAGCGCAGATAAAGTGGAG	1789
Db	1643	ATTTTCATACGCTGGCCCTACATCGGACACACTCCCTACGCTCGAGCTCGGGCG	1702
Qy	1790	GTCPTTCCAACTGTCTCTCTTCGAGGAATGGAAGCCATTCGAGAAAGCAATGCATG	1849
Db	1703	GGGTGTCCAAAGTGTCCTCTCGTTCCGCGGCACAACCCGGTACGCGAGGCATCCACT	1762
Qy	1850	GGGTTCCTTTACCAATGCAATCAAGTCTGGCATGGACATGGAGATAGTCAATGCTGGAA	1909
Db	1763	CGGTGTTCTCTACTACGCGATCGCAACGGCCTGACATGGGCATCGTCAACGCCGGCC	1822
Qy	1910	ACCTCCCTGTGTATGATGATATCATTAAGGAACTTTCGACGCTGTGTGAAGATCTCATCT	1969
Db	1823	AGCTGGAAATCTACGACGATTCGGAAGCGCTCGCGACCGGTTCGAGGACGTGGTGC	1882
Qy	1970	GGAATAAGACCTTGAGGCCACTGAGAAAGCTCTTAAGTTATGCCGACACTCAA---GGCA	2026
Db	1883	TCAACCGCAGCCCGAGGCCACGAGGCCCTGTGTGGCGATCGCGCAGCACTACAAGGGCG	1942
Qy	2027	CAGGAGGAAGAAGTCAATTCAGACTGATCAGTGGAGAAATGCCCCCTGTGCGAAGACGCC	2086
Db	1943	CGCGCGCGGTCAAGAGCGCGAGGACAGGAATGGCGCAGCTACGCTCGAAGAGGCC	2002
Qy	2087	TTGAGTATGCCCTGTGTAAGGCAATCGAAAAATATTTTGAAGGATCTAGAGGAAGCCA	2146
Db	2003	TCGAGCATGGCTGGTCAAGGSCATCACCTGGATCGTCGAGGACACCGAGGA---	2057
Qy	2147	GGTTAAACCAAAAAATATCCCGACCTCTCAATATAATGAAGGACCCCTGATGAATG	2206
Db	2058	----ATGCGCGCAGCACTGTGCGCGTCCCATCGAGGTCAATCGAAGGTCCGCTGATGTCCG	2113
Qy	2207	GAATGAAAAATGTTGGTGATCTTTTGGAGCTCGAAAAATGTTTCTACCTCAGGTTATAA	2266
Db	2114	GGATGAACGTGGTCGCGACCTGTGTGCGCCCGCAAGATGTTCTCTCCGCAAGTGTGCA	2173
Qy	2267	AGTCAGCCCCGGTTATCAAGAAAGCTGTGGCCACCTTATCCCTTTTCATCGAAAAAGAAA	2326
Db	2174	AGTCGCGCGAGTGATGAAGCAGGGCGTGGCCCACTGATTCTCTTCATCGAGCGGAGA	2233
Qy	2327	GAGAAGAAACGAGATGCTTTAAACGGCACAGTAGAAGAAAGGACCTTTACGAGGCAACA	2386
Db	2234	AAGCGACA-----AGCGGAAGCCAAAGGCAAGA	2263
Qy	2387	TCGTCTGGCCACTGTTTAAAGGCGAGCTGCACGACATAGCAGCAACATAGTTGGAGTAG	2446
Db	2264	TCCTGATGGCCACGGTGAAGGGCGAGTGCAACACATCGGCACAGACATCGTCGGCGTGG	2323
Qy	2447	TCCTTGGCTGCAATAATTTCCGAGTTATTCATTTTAGGAGTCATGACTCCCATGTGATAAGA	2506
Db	2324	TGCTCGGCTGCAACGGCTATGACGTGTGACCTCGCGTGATGGTGCCGCGCGAGNAGA	2383
Qy	2507	TACTGAAAGCTGCTCTTGACCAAAAGCAGATATAATTTGSCCTGTTCAGGACTCATCACTC	2566
Db	2384	TCCTGCAGACGCCATCGCGAGAAATGCGACATCATCGSCCTGTCTGGCCTGATCAAGC	2443
Qy	2567	CTTCCCTGGAATAATGATTTTTTGTGCCAAGGAATGAGAGATTAGCTATAAGGATTC	2626
Db	2444	CGTCGCTGGACAGATGGTCCACGTGCCCAAGGAAATGACGCGGCGAGAAATTTCCAGTTGC	2503
Qy	2627	CATTGTTGATTGGAGGAGCAACCACTTCAAAAAACCCACACAGCAGTGTAAAAATAGCTCCGA	2686
Db	2504	CGCTGATGATCGGCGGCCACTACTCTCGAAGGGGCATACCGGGGTGAAGTGCATCCGC	2563
Qy	2687	GATACAG---TGCACTCTGAATCCATGTCTCTGACGCGTCCAAAGAGTGTGGTGTGTGTT	2743
Db	2564	AGTACAGCAACGACGCGGTGCTCTACGTCAACCGACGCTCCGCGCGGTAGGGGTGGCCA	2623
Qy	2744	CCGAGCTGTTAGTAAATCTAAGAGTGAATATCTTTGAGGAAATCATCGGAAGAAATATG	2803
Db	2624	CCAGCTGCTGTCCAAAGAGCTGAAGCCGACTACGTGGCCCGCACCGCGCGGACTACG	2683







QY 2327 GAGAGAAACAGAGTGTCTTAACGGCACAGTAGAAGAGAGACCCCTTACAGGGCACCA 2386  
D 2334 AAGCGACA-----AGCCGAGAGCCAAAGGCAAGA 2263  
QY 2387 TCGTCTGGCCACTGTTAAAGGCGACGTGACGACATAGGCAAGAACATAGTTGGAGTAG 2446  
D 2264 TCCTGATGGCCACCGTGAAGGGCGACGTGACGACATCGGCAAGAACATCGTCGGCGTGG 2323  
QY 2447 TCCTTGGCTGCAATAATTTCCGAGTTATTTGATTTAGGAGTCACTACTCTCATGTGATAGA 2506  
D 2324 TGCTCGGCTGCAACGGCTATGACGTGTCGACCTCGGCGTGATTTGGTCCGCGCGAGAGA 2383  
QY 2507 TACTAAAGCTGCTCTTGACCAAAAGCAGATATAATTGGCTGTGACGACTCATCACTC 2566  
D 2384 TCCTGACAGCGCCATCGCGAGAAAATGCGACATCATCGGCTGTCTGGCCTGATCAGC 2443  
QY 2567 CTTCCCTGGATGAATGATTTTGTGTCACAGGAAAATGGAGAGATTAGCTATAGGATTC 2626  
D 2444 CGTCGCTGGACGAGATGGTCCACGTGCGCAAGGAAAATGACGCGCAGAAATTTCCAGTTGC 2503  
QY 2627 CATTTGTTGATTTGGAGGACCAACCTTCAAAAACCCACACAGCAGTTAAATAGTCCGA 2686  
D 2504 CGCTGATGATCGGCGCGCCACTACTCGAAGGCGCATACCGCGTGAGATCATCCGC 2563  
QY 2687 GATACAG---TGACACTGTATTCATGTCCTGGACGCGTCCAAAGAGTGTGGTGTGTT 2743  
D 2564 AGTACAGCAACGACGCGTGTGTCTAGGTCAACGACGCTCGCGCGTGGCGCA 2623  
QY 2744 CCCAGCTTTAGATGAATAATCTAAGGATGAATCTTTGAGGAATCATGAAAGATATG 2803  
D 2624 CCAGCTCTGTCTCAAGGAGCTGAAGCGCGACTAGTGGCGCGCACCGCCGCGACTACG 2683  
QY 2804 AAGATAT-----TAGACAGGACCAATTAATGAGTCTCTCAAGGAGAGGATACTTACCCCTT 2858  
D 2684 CGGTGGTCCGGAACGACGCGCAACCCGACGCGCCGCAACCGAGCGGCTGAGCTACGAA 2743  
QY 2859 AAGTCAAGCCAGAAAAGTGGTTTCCAAATGATTTGGCTGTCTGAACCTCACCCAGTGA 2918  
D 2744 AGGCGATCGCCAAACGACGCGGTTCGACTGGCGCGTACACAGGCGCGAC----- 2795  
QY 2919 GCCACGTTTATTTGGACCCAGGCTTTTGAAGACTATGACTGACGAAAGCTGGTGGACTA 2978  
D 2796 GCCTTCTTCCCGCGCTCAGGCTGCTCGACGAGATCGACCTCGCGGTGCTCGCGCGATA 2855  
QY 2979 CATTGACTGAAGCTTTCTTTGATGCTGGCAGCTCCGGGCGAAGTACCCGAATCGAGG 3038  
D 2856 CATGACTGGACGCGCTTCTTCAATTTCTTGGACCTTGGCCGCGCAAGTACCCG----- 2907  
QY 3039 CTTCCCAAGATATTTAAGCAGCAAAACAGTAGTGGAGTGGAGGCGCAGGAAGGTCTACGATGA 3098  
D 2908 -----CGCATCTCACCGACGAGGTGTCGCGAGGCGCGCACCTCGTTGTTCAAGCA 2960  
QY 3099 TGCCCACAATATGTGAAACACTGATTTAGTCAAAGAAACTCCGGGCGCGGGGTGGGT 3158  
D 2961 GCCCAGGCGATGCTGAAGAGCTGATCGACGAGAAGCTGATCAAGGCGCGCGCGGTGT 3020  
QY 3159 TGGGTTCTGGCCAGCAGAGATATCCAAGACGACATTCACCTGTACGAGGCTGCTGT 3218  
D 3021 CGGCTTCTGGCGGCGCAACCAAGTTCGACGACACACCTGAGAGGTCTACGCGCGCGATGG 3080  
QY 3219 GCCCCAGGCTGCAGAGCCCATAGCCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGA 3278  
D 3081 C-----GAGACCTCGCACCTTGACACCTGCGGCGAGCAGATCAAGCC 3128  
QY 3279 CTCTGCCAGCAGGACCATACTGCTCTCAGACTTCATCGCTCCCTCTGCAATCTGG 3338  
D 3129 GAGCGCA-----AGCCGACCTCTGCTGCGCCGATTTCTGCGCGCGAAGAAAGCG 3182  
QY 3339 CATCGTACTACTCTGGGCTGTTG-----CGTTGCTGCTTTGGGTAGAGAGCTAG 3395  
D 3183 CGTGGCGCATCATCGCGGCTTTCATCACCCCGCGGATTCGCGCGCGAGGAAGTGGC 3242  
QY 3396 CAAGGCTTATGAGGATGATGGTGACGACTATACAGCAGCATCATGCTGTCAGGCGCTGGGGA 3455

D 3243 CAAGGCGTACGAAGCCAAAGGCGACGACTACAACGACATCATGTTCAAGCGCTCGCCGA 3302  
QY 3456 CCAGCTGGCAGAGAGCCCTTTTCAGAAAGAGCTTCCATGAAGAGTTTCGCGAGAACTGTGGGC 3515  
D 3303 CCAGCTGGCAGAGAGCCCTTCGCGCGAGTGGCTGCGAGCGGGTGGCAGAGTACTGGGG 3362  
QY 3516 CTACTGTGGCAGTGAAGCAGCTGACGAGCTGCGAGACCTGCGAAGGTTGCGGTACAAGGCGAT 3575  
D 3363 CTAGCCCGCGAGCAACACCTCGACACGAGGCTTTGATCAAGGAGCAATACGTCGCGCAT 3422  
QY 3576 CCGCCCGGCTCTGGCTTACCCCGAGCCGCGAGCCACCGAGAGAGCTCACCATGTGGAG 3635  
D 3423 CCGCCCGGCTACCGGGCTTACCGCGCTGCGCCCGACCATACCGAGAAAGGCACTCTCTTTCGA 3482  
QY 3636 ACTCGCAGACATCGACGAGCTACAGGCACTTAGGTTAAAGATCATTTAGCAATGGCAC 3695  
D 3483 ACTGCTCGATCCGAGGCGCTGTCGCGCGTACGCTGACCGACACTACCGGATGTTCCC 3542  
QY 3696 TGCTTCAGCAGTCTCAGGCGCTTACTTCTCAATTTTGAAGTCCAAATATTTTGTGTGGG 3755  
D 3543 GCGCGCGCGTTCAGCGTTGGTATTTTCGCCCCACCGCGCGCAGTACTTCGCGGTGG 3602  
QY 3756 GAGATTTCCAAGATCAGGTTGAGGATTTATGCAATTGAGGAAGAACATATCTGTGGCTGA 3815  
D 3603 CAAGATCGACAAGGACGAGGTGGAAACGCTACAGCGCAAGGCGCGAGGAAGCCAGCGT 3662  
QY 3816 GGTTCAGAAATGGCTTCGACCCATTTTGGGATATGAT 3852  
D 3663 CAGCGAGGCTGGCTGGCGCGCACTTGGCTACGAT 3699

## RESULT 15

ADP02792  
ID ADP02792 standard; DNA; 3750 BP.

XX ADP02792;

XX AC

XX AC

DT 12-FEB-2004 (first entry)

XX Bacterial polynucleotide #3077.

DE Proteus mirabilis infection; bacterial infection; antibacterial;  
XX immunostimulant; gene; ds.

KW Proteus mirabilis.

XX Proteus mirabilis.

OS US6605709-B1.

XX 12-AUG-2003.

PD 05-APR-2000; 2000US-00543681.

XX 09-APR-1999; 99US-0128706P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PA Breton GL;

XX WPI; 2003-895291/82.

XX P-PSDB; ADP06964.

DR New Proteus mirabilis polypeptides and polynucleotides, useful as

XX reagents for diagnosis of bacterial disease, as components of

PT antibacterial vaccines, as targets for antibacterial drugs, or as

XX biocontrol agents for plants.

XX Disclosure; SEQ ID NO 3077; 870pp; English.

XX The invention relates to new Proteus mirabilis polypeptides and

XX polynucleotides. The invention also relates to antibodies against the

CC polypeptides, methods for producing the polypeptides, a method of

CC generating vaccines for immunising an individual against P. mirabilis, a





Qy 2116 AACATATTAATTGAGGATCTAGAGAGCCAGGTTAAACCAAAATAATATCCCGACCT 2175  
 Db 2086 GAATTTATCAATTGAATACAGAAG-----CGTCCGCCCAACAAGCTTCAGTCTCT 2136  
 Qy 2176 CTCAATATAATGAGGAGCCCTGATGATGAAATGAAATGTTGGTGATCTTTTGGGA 2235  
 Db 2137 ATCGAAGTGATAGAGGGCCATTAATGAATGGCATGAACACAGTGGGCGATCTCTTTGGT 2196  
 Qy 2236 GCTGGAAATGTTTCTACCTCAGGTTATAAGTCAGCCCGGGTTATGAAAGCGCTGTT 2295  
 Db 2197 GAAGGTAAATGTTTACCCCAAGTAGTAAATCGGCACGGGTGATGAACAGGCCGTA 2256  
 Qy 2296 GGCACACCTTATCCCTTTTCATCGAAAGAAAGAGAGAAACACAGAGTGCTTAAACGGCACA 2355  
 Db 2257 GCTTACCTTGAACTTATTAATCAAGCGACAAAGACAGCGAG----- 2300  
 Qy 2356 GTAGAAGAGGAGCCCTTACAGGGCACCAATCGTGTGGCCTGTTTAAAGGCGAGTG 2415  
 Db 2301 -----CTCGCGAGAAAGTGTTTACTGGCGACAGTAAAGGAGAGATGTC 2343  
 Qy 2416 CACGACATAGCAAGACATAGTGGAGTAGTCTCTGGCTGCAATATTTCCGAGTTATT 2475  
 Db 2344 CACGACATGGCAAAATATTTGTCCGAGTGTTGCAATGTAATTAATGAAATATTC 2403  
 Qy 2476 GATTTAGAGTCATGACTCCATGTGATAAGATACCTGAAAGCTGCTCTTGACCAACAAGCA 2535  
 Db 2404 GATCTGGGGTCATGCTGATCTCTGCGATGATGTTGCAACAGCAATGAGCAAAATGTT 2463  
 Qy 2536 GATTAATTTGGCTGTGAGGACTCATCTCTCTCCCTGGATGAAATGTTTTGTTGCC 2595  
 Db 2464 GATATTTAGGATTTGCTGGATGATGATCACTCACTGGATGAGTGGTGAATGTTGCT 2523  
 Qy 2596 AAGGAATGAGAGATGATGCTATAAGGATTTCCATTTGATTTGAGGAGGCAACACTTCA 2655  
 Db 2524 AAGGAGATGGAAGAAGCGCGTTTTCTCTCTCTTAAATGATTTGGTGGCGACAACTCG 2593  
 Qy 2656 AAAACCCACACAGAGATTTAAATAGCTCCGAGATACAGTGACCTGTAATCCATGTCCTG 2715  
 Db 2584 AAAGGCACACGGCAGTAAATTTGAACCTAATTAACAGCCACCAACCGTTTATGTGAG 2643  
 Qy 2716 GACCGCTCAAAGAGTGGTGGTGTGTTCCAGCTGTTAGATGAAATCTAAAGGATGAA 2775  
 Db 2644 AATGCGTCAAGAAACGGTCCGGGTGCTGCTGCTCAATTTATTCGGCCACACAAAGGCTGAT 2703  
 Qy 2776 TACTTTGAGGAATCATGGAGAAATATGAAGATTTAGACAGGACCAATATGAGTCTCTC 2835  
 Db 2704 TTTGCTGCTAAACAACGCGCGTGAATATGAAGTGGTACGTGAGCAGTATGCGAGAAAAA 2763  
 Qy 2836 AAGGAGAGGAGATACCTTACCTTAAAGTCAAGCCAGAAAAAGTGGTTTCCAAATGGATTGG 2895  
 Db 2764 CCGGTACGCTCCGTTTCTTTAGCGACAGCTCGTGCCATGCAATTCGACCTTGACTGG 2823  
 Qy 2896 CTGTCTGAACCTCAACCCAGTGAAGCCCAAGTTTATGGGACCCAGGCTTTTGAAGACTAT 2955  
 Db 2824 CAACACTATAACCCACCTTAAGCCGAACCAATTTAGGGGTACAACAAGTTTACAGCCAAATTT 2883  
 Qy 2956 GACCTGCAAGAGCTGGTGACTCAATTTGACTGGAAGCCCTTTCTTTGATGTCCTGGCAGCTC 3015  
 Db 2884 GAA-----ACTTTACGTGAGTATATCGACTGGAGCGCTTTCTTTATGACTTGGTTCATTG 2937  
 Qy 3016 CGGGGAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGCACAACAGTAGGTGGA 3075  
 Db 2938 CGGGGAAATATCCACGT-----ATAATTAGAGATGAGGTTGTCCGAGAA 2982  
 Qy 3076 GAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTGATTTAGTCAAAAG 3135  
 Db 2983 GAGGCAAGACGCGTTTTTGGCAGATGCGAAATGCCATGTTGGATAACTCAGCCGTGAAAAA 3042  
 Qy 3136 AAATCTCGGGCCCGGGGTGGTGGGTTCTGGCCAGCAGAGATATCCAAAGACGACATT 3195  
 Db 3043 TTACTGACACCGAAAGGTATCGTTGGTTTATTCCTCGCTAATTCGCTTGGTGTGATGACATT 3102

Qy 3196 CACCTGTATCAGAGAGGCTGTGTGCCAGGCTGCAGAGCCCATAGCACCCTTTCTATGGG 3255  
 Db 3103 ATTATTTTACCAAGATGAAACGCGACAGCATG-----AGTTATTACGTTGTTGCCAT 3153  
 Qy 3256 TTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCAGGAGCCATCTACTGCTCTCAGAC 3315  
 Db 3154 TTACGCCCAACAAACAGAAAAA-----GAGTACCCCTAACTATTGCTTTAGCAGAC 3204  
 Qy 3316 TTCAATCGCTCCCTTGGCAATCTGGCATCCGTGACTTACCTGGGCTGTTTGGCGTTGCTGC 3375  
 Db 3205 TTTATGCCCCCTGTGGATAGTGTCTTGCTGACTATTTTGGTGTCTTTGAGTTACCGGC 3264  
 Qy 3376 TTTGGGTAGAAGA-----GCTGAGCAAGGCCCTATAGGATGATGTTGACGACTACAGCAGC 3432  
 Db 3265 GGTCTAGAAGAGATGACCTCGCTAATGGGTATGATAATCTCATGATGATTACAAATAA 3324  
 Qy 3433 ATCATGTTCAAGGGCTGGGGGACCGGCTGGCAGAGCCCTTTGACAGAAAGACTCCATGAA 3492  
 Db 3325 ATTATGGTAAAGCGCTATCCGATCGTCTTGCAGAGCTTTTGCAGAAATATTTGATCA 3384  
 Qy 3493 AGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGAGCGTCGACAGCTG 3552  
 Db 3385 CAAAGTGAGAACCAAAATTTGGGGTTACAGTCTTGACGAGAGCGCTCTCTAATGACGAGTTG 3444  
 Qy 3553 CGAAGGTTGGCGTACAAGGGCATCCGCCCGGCTCTGGCTACCCAGCCAGCCGACAC 3612  
 Db 3445 ATTAGAGAAAAATATCAGGAAACACGACCTGACCCGGGATATCCAGCTTCCCGCAGAGC 3504  
 Qy 3613 ACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGCAATTAGGTTA 3672  
 Db 3505 ACAGAAAAAGCAAAATTTGGCAACTGTTAAATGTAGAAAAATCGCATTTGGTATGAAGCTT 3564  
 Qy 3673 ACAGAACTATTAGCAATGGCACCTGCTTACAGCAGTCTCAGGCTCTACTTCTCCAATTTG 3732  
 Db 3565 ACTGATGCTTATGCGATGTGGCGGGGCTTTCGGTATCGGGTTGGTATTTTAGTTCATCCT 3624  
 Qy 3733 AAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAGGATCAGGTTGAGGATTTATGCAATTG 3792  
 Db 3625 GAAAGTAAATATTTTGGCGTCTCAATTTCAAAGATCAGGTAGAAGATTAGCCCAA 3684  
 Qy 3793 AGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTTTGGGATATGA 3851  
 Db 3685 CGTCGTGATGAGTATGAGTGAAGTTGAGCGTTGGTTAGCCCCCTAATTTAGGTTATGA 3743

Search completed: March 6, 2006, 16:12:40  
 Job time : 1510 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 14:48:07 ; Search time 9723 Seconds  
(without alignments)  
18858.231 Million cell updates/sec

Title: US-10-607-712-1  
Perfect score: 3919  
Sequence: 1 ggcacctgtggagacgcg.....ctcaaggaaatacaacctag 3919

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_hc.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_est7.\*  
9: gb\_gss1.\*  
10: gb\_gss2.\*  
11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3632.4	92.7	3711	11	DQ037675
2	2977.6	76.0	3677	11	DQ037676
3	2282.6	58.2	4888	4	CR859882
4	1435.6	36.6	2150	4	AK037599
5	819.2	20.9	1788	4	AK085987
6	814.8	20.8	863	3	BM450887
7	814	20.8	1043	3	BQ062840
8	797.4	20.3	871	5	BX390174
9	791.2	20.2	943	5	BU189680
10	718.4	18.3	961	5	BX332266
11	707.4	18.1	721	1	AL598311
12	672.4	17.2	698	7	CN414972
13	611.4	15.6	715	8	DN280822
14	598.8	15.3	758	8	CX228713
15	591.2	15.1	623	7	CH629973
16	591.2	15.1	626	5	BX485695
17	578.8	14.8	582	3	BP336015
18	578.8	14.8	582	3	BP336553
19	570.4	14.6	583	3	BP313585
20	570.4	14.6	583	3	BP337383
21	568.8	14.5	997	2	BG258321
22	561.2	14.3	618	7	CV356711

23	558.2	14.2	915	5	BQ954233
24	555.2	14.2	560	3	BP220649
25	554.6	14.2	688	1	AI872333
26	549.2	14.0	806	1	AJ724455
27	537.8	13.7	581	3	BP336596
28	534.2	13.6	937	5	BQ960377
29	531.2	13.6	657	9	CC249505
30	530.2	13.5	874	2	BE879762
31	527.8	13.5	727	6	CB600441
32	527.2	13.5	757	1	AJ724469
33	520.6	13.3	527	5	BX470290
34	517.8	13.2	574	7	CV344680
35	505	12.9	617	7	CK620511
36	503	12.8	581	8	DN279755
37	502.4	12.8	658	6	CB455191
38	501.8	12.8	885	6	CF548421
39	490	12.5	490	1	AL041796
40	486.4	12.4	914	5	BU343179
41	481.4	12.3	665	1	BB156044
42	479.2	12.2	517	3	BQ311389
43	478.4	12.2	496	5	BU429497
44	475.6	12.1	717	1	AJ724468
45	475.4	12.1	717	6	CF739229

## ALIGNMENTS

RESULT 1  
LOCUS DQ037675 3711 bp DNA linear GSS 02-JUN-2005  
DEFINITION Homo sapiens MTR gene, VIRTUAL TRANSCRIPT, partial sequence.  
ACCESSION DQ037675  
VERSION DQ037675.1 GI:66888884  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3711)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees  
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)  
PUBMED 15869325  
REFERENCE 2 (bases 1 to 3711)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.  
FEATURES  
source  
1..3711  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
<1..>3711  
/gene="MTR"  
/locus\_tag="HC7587"

ORIGIN  
Query Match 92.7%; Score 3632.4; DB 11; Length 3711;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 3636; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy	64	ATGTCACCGCGCTCCAAAGACCTGTGCAACCCGAAAGTCTGAAGAAACCCCTGCGGAT	123
Ds	1	ATGTCACCGCGCTCCAAAGACCTGTGCAACCCGAAAGTCTGAAGAAACCCCTGCGGAT	60
Qy	124	GAGATCAATGCCATTCTGCAGAGAGGATTATGCTGCTGGATGAGGATGAGGACCATG	183
Ds	61	GAGATCAATGCCATTCTGCAGAGAGGATTATGCTGCTGGATGAGGATGAGGACCATG	120
Qy	184	ATCAGCGGAGAGCTTAAACGAAAGACATCTCCGAGGTGAGGATTTAAAGATCATGCC	243
Ds	121	ATCAGCGGAGAGCTTAAACGAAAGACATCTCCGAGGTGAGGATTTAAAGATCATGCC	180
Qy	244	AGGCGCTGAAAGGCAACAATGACATTTTAAGTATACTCAGCCCTGATGTCATTTACCA	303
Ds	181	AGGCGCTGAAAGGCAACAATGACATTTTAAGTATACTCAGCCCTGATGTCATTTACCA	240
Qy	304	ATCCATAAGGAATACTTGTGCTGGCTGGGCGAGATATCATTTGAAACAAATACTTTAGCAGC	363
Ds	241	ATCCATAAGGAATACTTGTGCTGGCTGGGCGAGATATCATTTGAAACAAATACTTTAGCAGC	300
Qy	364	ACTAGTATGCGCAAGCTGACTATGCGCTTGAAACACTTGGCTACCGATGAAATGTC	423
Ds	301	ACTAGTATGCGCAAGCTGACTATGCGCTTGAAACACTTGGCTACCGATGAAATGTC	360
Qy	424	TCTCAGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGAGG	483
Ds	361	TCTCAGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGAGG	420
Qy	484	TTTGTGGCAGGGGCTCTGGGCTCCGACTAATAAGACACTCTCTGTGTCGCCATCTGTGGAA	543
Ds	421	TTTGTGGCAGGGGCTCTGGGCTCCGACTAATAAGACACTCTCTGTGTCGCCATCTGTGGAA	480
Qy	544	AGGCGGAGTATAGGAAACATCACATTTGATGAGCTTGTGAGGATACCAAGAGAGGCC	603
Ds	481	AGGCGGAGTATAGGAAACATCACATTTGATGAGCTTGTGAGGATACCAAGAGAGGCC	540
Qy	604	AAAGGACTCTGGAATGGCGGGTGGATATCTTACTCATTTGAAATCTATTTTGTGACTGCC	663
Ds	541	AAAGGACTCTGGAATGGCGGGTGGATATCTTACTCATTTGAAATCTATTTTGTGACTGCC	600
Qy	664	AATGCCAAGGAGCGCTTGTGCTCCAAATCTTTTGGAGGAAATATGCTCCCGG	723
Ds	601	AATGCCAAGGAGCGCTTGTGCTCCAAATCTTTTGGAGGAAATATGCTCCCGG	660
Qy	724	CCTATCTTTATTCAGGACGATGTTGATTAAGTGGCGGACTCTTTCGGGACAGACA	783
Ds	661	CCTATCTTTATTCAGGACGATGTTGATTAAGTGGCGGACTCTTTCGGGACAGACA	720
Qy	784	GGAGAGGATTTGTATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAATGTT	843
Ds	721	GGAGAGGATTTGTATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAATGTT	780
Qy	844	GCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATAATTTGGAATAATGACACGCC	903
Ds	781	GCTTTGGGTGAGCTGGAATGAGACCTTTTATTTGAAATAATTTGGAATAATGACACGCC	840
Qy	904	TATGTCCTCTGTTATCCCAATGCGAGCTTTCCTCAACACTTTTGGTGTGACTATGATGAACG	963
Ds	841	TATGTCCTCTGTTATCCCAATGCGAGCTTTCCTCAACACTTTTGGTGTGACTATGATGAACG	900
Qy	964	CTTCTATGATGGCCCAAGCCTAAAGGATTTTGTCTATGGATGGCTTGGTCAATATAGTT	1023
Ds	901	CTTCTATGATGGCCCAAGCCTAAAGGATTTTGTCTATGGATGGCTTGGTCAATATAGTT	960
Qy	1024	GGAGGATGCTGTGGGTCAACACGATCATATCAGGGAATTTGCTGAAAGCTGTGAAAAAT	1083
Ds	961	NN	1020
Qy	1084	TGTAAGCCTAGAGTTCCACCTGCCATGCTTTTGAAGGACATATGTTACTGTCTGGTCTA	1143
Ds	1021	TGTAAGCCTAGAGTTCCACCTGCCATGCTTTTGAAGGACATATGTTACTGTCTGGTCTA	1080

Qy	1144	GAGCCCTTCAGGATTCGACCGTACACCAACTTTGTTAAACATTCGAGAGCGCTGTAATGTT	1203
Ds	1081	GAGCCCTTCAGGATTCGACCGTACACCAACTTTGTTAAACATTCGAGAGCGCTGTAATGTT	1140
Qy	1204	GCAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAACTATGAGAGCGCTTGTGT	1263
Ds	1141	GCAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAACTATGAGAGCGCTTGTGT	1200
Qy	1264	TTTGGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCACCATGATGATGGC	1323
Ds	1201	TTTGGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCACCATGATGATGGC	1260
Qy	1324	ATGCTAGATCCAGTGCATGACCCAGATTTTGCACCTTAATTTGCTCCGAGCCAGAC	1383
Ds	1261	ATGCTAGATCCAGTGCATGACCCAGATTTTGCACCTTAATTTGCTCCGAGCCAGAC	1320
Qy	1384	ATGCGAAAGTACTTTTGTGCATCGACTCTCCAAATTTTGTCTGTGATGAGCTGGGTTA	1443
Ds	1321	ATGCGAAAGTACTTTTGTGTGCATCGACTCTCCAAATTTTGTCTGTGATGAGCTGGGTTA	1380
Qy	1444	AAGTGTGCGCAAGGAAGTGCATTTGCAATAGCATTAGTCTGAAGGAAGGAGAGCGAC	1503
Ds	1381	AAGTGTGCGCAAGGAAGTGCATTTGCAATAGCATTAGTCTGAAGGAAGGAGAGCGAC	1440
Qy	1504	TTCTTGGAGAGGCCAGGAAGATTTAAAGATATGGAGCTGCTATGTTGTTTTCATGGCTTTT	1563
Ds	1441	TTCTTGGAGAGGCCAGGAAGATTTAAAGATATGGAGCTGCTATGTTGTTTTCATGGCTTTT	1500
Qy	1564	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGCTTAC	1623
Ds	1501	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGGCTTAC	1560
Qy	1624	CATCTGCTTCTGAAAAAATCGGGCTTTAAATCCAAAATGACATTAATTTTGAACCTTAATTC	1683
Ds	1561	CATCTGCTTCTGAAAAAATCGGGCTTTAAATCCAAAATGACATTAATTTTGAACCTTAATTC	1620
Qy	1684	CTAACCATTCGGGACTGGAAATGGAGGAAACACAACCTTGTATGCCATTAATTTTATCCATGCA	1743
Ds	1621	CTAACCATTCGGGACTGGAAATGGAGGAAACACAACCTTGTATGCCATTAATTTTATCCATGCA	1680
Qy	1744	ACAAAGTCAATTAAGAAAACATTTACCTGGAGCGCAAGATTAAGTGGAGGTCTTTTCAACTTG	1803
Ds	1681	ACAAAGTCAATTAAGAAAACATTTACCTGGAGCGCAAGATTAAGTGGAGGTCTTTTCAACTTG	1740
Qy	1804	TCCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGTTTCTTCTTAC	1863
Ds	1741	TCCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGTTTCTTCTTAC	1800
Qy	1864	CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTTCCCTGTGTAT	1923
Ds	1801	CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTTCCCTGTGTAT	1860
Qy	1924	GATGATATCCATAAGGAACCTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCT	1983
Ds	1861	GATGATATCCATAAGGAACCTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCT	1920
Qy	1984	GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGACAGAGGAGGAGAAAGTC	2043
Ds	1921	GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGACAGAGGAGGAGAAAGTC	1980
Qy	2044	ATTCAGACTGATAGTGGAGAAATGGCCCTGTCCAGAAAGCCCTTGAAGTATGCCCTTGTG	2103
Ds	1981	ATTCAGACTGATAGTGGAGAAATGGCCCTGTCCAGAAAGCCCTTGAAGTATGCCCTTGTG	2040
Qy	2104	AAGGCCATTGAAAAACATATTTATTGAGGATCTAGGAGAGCCAGGTTTAAACAAAAAAA	2163
Ds	2041	AAGGCCATTGAAAAACATATTTATTGAGGATCTAGGAGAGCCAGGTTTAAACAAAAAAA	2100
Qy	2164	TATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGATGAAATGTTGTTGTT	2223
Ds	2101	TATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGATGAAATGTTGTTGTT	2160
Qy	2224	GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATTAAGTCAAGCCCGGGTTATG	2283

```
Db 2161 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCGGGTTATG 2220
Qy 2284 AAGAAGGCTGTTGGCCACCTTATCCCTTTCATGAAAAAGAGAAAGAACAGAGTG 2343
Db 2221 AAGAAGGCTGTTGGCCACCTTATCCCTTTCATGAAAAAGAGAAAGAACAGAGTG 2280
Qy 2344 CTTAAACGGCAGTAGAAGAGAGAGCCCTTACCAGGGCCACCTGCTGGCCACCTGTT 2403
Db 2281 CTTAAACGGCAGTAGAAGAGAGAGCCCTTACCAGGGCCACCTGCTGGCCACCTGTT 2340
Qy 2404 AAAGCGGACGTGACGACATAGGCAAGAAACATAGTTGGAGTAGTCCCTTGGCTGCAATAAT 2463
Db 2341 AAAGCGGACGTGACGACATAGGCAAGAAACATAGTTGGAGTAGTCCCTTGGCTGCAATAAT 2400
Qy 2464 TTCCGAGTTATGATTAGGAGTCAATGCTCCATGATAGATACATGAAAGCTGCTCTT 2523
Db 2401 TTCCGAGTTATGATTAGGAGTCAATGCTCCATGATAGATACATGAAAGCTGCTCTT 2460
Qy 2524 GACCACAAGCAGATATAATTTGGCTGTCCAGGACTCATCTCTTCCCTGGATGAATG 2583
Db 2461 GACCACAAGCAGATATAATTTGGCTGTCCAGGACTCATCTCTTCCCTGGATGAATG 2520
Qy 2584 ATTTTGTGTGCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTGTTGATGGAGGA 2643
Db 2521 ATTTTGTGTGCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTGTTGATGGAGGA 2580
Qy 2644 GCACACCTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTGTA 2703
Db 2581 GCACACCTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTGTA 2640
Qy 2704 ATCCATGCTCTGAGCGGTCCAGAGTGTGTGTGTGTTCCAGCTGTAGATGAATAAT 2763
Db 2641 ATCCATGCTCTGAGCGGTCCAGAGTGTGTGTGTGTTCCAGCTGTAGATGAATAAT 2700
Qy 2764 CTAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATAGACAGGACCAT 2823
Db 2701 CTAAGGATGAATACTTTGAGGAATCATGGAAGATATGAAGATATAGACAGGACCAT 2760
Qy 2824 TATGAGTCTCTCAAGGAGAGAGATCTTACCTTTAAGTCAAGCCAGAAAAAGTGTTC 2883
Db 2761 TATGAGTCTCTCAAGGAGAGAGATCTTACCTTTAAGTCAAGCCAGAAAAAGTGTTC 2820
Qy 2884 CAATGAGATTGGCTGTGAACCTCACCCAGTGAAGCCACGTTTATGGACCCAGGTC 2943
Db 2821 CAATGAGATTGGCTGTGTGAACCTCACCCAGTGAAGCCACGTTTATGGACCCAGGTC 2880
Qy 2944 TTTCAGACTATGACTGCGAGAGCTGTGTGACTACATGACTGGAAGCTTTCTTTGAT 3003
Db 2881 TTTCAGACTATGACTGCGAGAGCTGTGTGACTACATGACTGGAAGCTTTCTTTGAT 2940
Qy 3004 GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAACGACAA 3063
Db 2941 GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAACGACAA 3000
Qy 3064 ACAGTAGGTGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTG 3123
Db 3001 ACAGTAGGTGAGAGGCCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTG 3060
Qy 3124 ATTAGTCAAAAGAACTCCGGGCCGGGGTGTGTGGTCTGGCCAGCAGAGATATC 3183
Db 3061 ATTAGTCAAAAGAACTCCGGGCCGGGGTGTGTGGTCTGGCCAGCAGAGATATC 3120
Qy 3184 CAAGCAGCATTTACCTGTATGAGAGGCTGTGTGCTGCCAGGCTGAGAGCCCATAGCC 3243
Db 3121 CAAGCAGCATTTACCTGTATGAGAGGCTGTGTGCTGCCAGGCTGAGAGCCCATAGCC 3180
Qy 3244 ACTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCGAGCAGGAGCCATATAC 3303
Db 3181 ACTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCGAGCAGGAGCCATATAC 3240
Qy 3304 TGCTCTCAGACTTTCATGCTCCCTTGCATTTGGCATTCGGTGAATACCTGGGCGCTGTTT 3363
```

```
Db 3241 TGCTCTCAGACTTTCATGCTCCCTTGCATTTGGCATTCGGTACTACCTGGGCGCTGTTT 3300
Qy 3364 GCCGTTGCTGCTTTGGGAGTAGAGAGCTGAGCAAGGCTTATGAGATGATGTGACGAC 3423
Db 3301 GCCGTTGCTGCTTTGGGAGTAGAGAGCTGAGCAAGGCTTATGAGATGATGTGACGAC 3360
Qy 3424 TACAGCAGCATCATGCTCAAGGCGCTGGGGGACCGGCTGCGAGAGGCTTTTGAGAGAG 3483
Db 3361 TACAGCAGCATCATGCTCAAGGCGCTGGGGGACCGGCTGCGAGAGGCTTTTGAGAGAG 3420
Qy 3484 CTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGGAGCTC 3543
Db 3421 CTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGGAGCTC 3480
Qy 3544 GCAGACTGCGAAGGTTGCGGTACAAGGGCATCCGCGGCTCTCGCTACCCAGCCAG 3603
Db 3481 GCAGACTGCGAAGGTTGCGGTACAAGGGCATCCGCGGCTCTCGCTACCCAGCCAG 3540
Qy 3604 CCCGACACACCCAGAGGCTCACCATGTGGAGACTCGCAGACATCAGAGCAGTCTACAGGC 3663
Db 3541 CCCGACACACCCAGAGGCTCACCATGTGGAGACTCGCAGACATCAGAGCAGTCTACAGGC 3600
Qy 3664 ATTAGGTTAACAGAACTATTAGCAATGGCACTGCTTACAGCAGTCTCAGGCTCTACTTC 3723
Db 3601 ATTAGGTTAACAGAACTATTAGCAATGGCACTGCTTACAGCAGTCTCAGGCTCTACTTC 3660
Qy 3724 TCCAATTTGAAGTCCAATAATTTTGTGTGGGGAAGATTTCCAAGGATCAG 3774
Db 3661 TCCAATTTGAAGTCCAATAATTTTGTGTGGGGAAGATTTCCAAGGATCAG 3711
```

## RESULT 2

DQ037676

LOCUS

DEFINITION

Pan troglodytes MTR gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION

DQ037676

VERSION

DQ037676.1

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

REFERENCE

AUTHORS

Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE

A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees

JOURNAL

(er) PLoS Biol. 3 (6), E170 (2005)

PUBMED

15869325

REFERENCE

AUTHORS

Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

TITLE

Direct Submission

JOURNAL

Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES

source

1..3677

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

gene

&lt;1..3677

/genes="MTR"

/locus\_tag="HC7587"

ORIGIN

Query Match

76.0%; Score 2977.6; DB 11; Length 3677;



Best Local Similarity 81.3%; Pred. No. 0;  
Matches 2989; Conservative 0; Mismatches 688; Indels 0; Gaps 0;

Qy	98	AAGTCTGAAAGAAACCCCTGCGGATGAGATCAATGCCATCTCAGAGAGGATTATGG	157
Db	1	AAGGTCTGAAAGAAACCCCTGCGGATGAGATCGATGCCATCTCAGAGAGGATTATGG	60
Qy	158	TGCTGGATGGAGGATGGGACCAATGATCCAGCGGAGAGCTAAACGAAGAACACTTCC	217
Db	61	TGCTGGATGGAGGATGGGACCAATGATCCAGCGGAGAGCTAAAGGAGAACACTTCC	120
Qy	218	GAGTTCAGGAAATTAAGATCATGCCAGCGCGCTGAAAGGCAACAATGACATTTAAGTA	277
Db	121	GAGTTCAGGAAATTAAGATCATGCCAGCGCGCTGAAAGGCAACAATGACATTTAAGTA	180
Qy	278	TAACTCAGCCGTGATGATTTACCAATCCATAGGAATATCTGCTGGCTGGGACGATA	337
Db	181	TAACTCAGCCGTGATGATTTACCAATCCATAGGAATATCTGCTGGCTGGGACGATA	240
Qy	338	TCATTGAAACAAATACTTTTAGCAGCACTAGTATTTGCCAAGCTGACTATGGCCTTGAAC	397
Db	241	TCATTGAAACAAATACTTTTAGCAGCACTAGTATTTGCCAAGCTGACTATGGCCTTGAAC	300
Qy	398	ACTTGGCCTACCGGATGAACATGTGCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGG	457
Db	301	ACTTGGCCTACCGGATGAACATGTGCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGG	360
Qy	458	TAACTCTCAGACAGGAATTAAGAGTTTGTGGCAGCGGGCTCTGGTCCGACTAATAGA	517
Db	361	TAACTCTCAGACAGGAATTAAGAGTTTGTGGCAGCGGGCTCTGGTCCGACTAATAGA	420
Qy	518	CACCTCTGTCTCCCATCTGTGAAAGCGCGGATTAAGAAATCATCATCTTCATGAGC	577
Db	421	CACCTCTGTCTCCCATCTGTGAAAGCGCGGATTAAGAAATCATCATCTTCATGAGC	480
Qy	578	TGTTTGAAGCATACCAAGACAGGCGCAAGACTTCTGGATGGGGGTTTGCATCTTAC	637
Db	481	TGTTTGAAGCATACCAAGACAGGCGCAAGACTTCTGGATGGGGGTTTGCATCTTAC	540
Qy	638	TCATTGAAACTATTTTGTGATCTGCCAATGCCAAGCGAGCGCTTGTGCACTCCAAATC	697
Db	541	TCATTGAAACTATTTTGTGATCTGCCAATGCCAAGCGAGCGCTTGTGCACTCCAAATC	600
Qy	698	TTTTTGAGGAGAAATATGCTCCCCGCTATCTTTATTTACGGACGATCGTGTATAAA	757
Db	601	NN	660
Qy	758	GTGGCGGACCTCTTCCGACAGACAGGAGGAGTTTGTCTCAGCGGTCTCATGGAG	817
Db	661	GTGGCGGACCTCTTCCGACAGACAGGAGGAGTTTGTCTCAGCGGTCTCATGGAG	720
Qy	818	AACCACTCTGCATTCGATTAATGTGCTTTGGGTGCAGCTGAGATGAGACTTTTATTG	877
Db	721	AACCACTCTGCATTCGATTAATGTGCTTTGGGTGCAGCTGAAATGAGACTTTTATTG	780
Qy	878	AAATAATTGGAATAATGTAACAGAGCTATGCTCTGTATCCCAATCGAGCTTCTCCA	937
Db	781	AAATAATTGGAATAATGTAACAGAGCTATGCTCTGTATCCCAATCGAGNNNNNNNN	840
Qy	938	ACACCTTTGGTGACTATGATGAAACGCTTCTATGATGCCCAAGCACCTAAAGGATTTG	997
Db	841	NN	900
Qy	998	CTATGGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAAACACAGATCATATCA	1057
Db	901	NN	960
Qy	1058	GGGAAATTTGCTGAAGCTGTGAAATAATTTGTAAGCTTAGAGTTCCACTGCCACTGTTTG	1117
Db	961	NN	1020
Qy	1118	AAGGACATATGTTACTGTCTGTCTAGAGCCCTTCCAGGATTCGACCGCTACCAACTTTG	1177
Db			

Db	1021	NN	1080
Qy	1178	TTAAACATTGGAGAGCGCTGTAATGTTGCAAGATCAAGAGAGTTTGTCTAAACTCATATGG	1237
Db	1081	TTAAACATTGGAGAGCGCTGTAATGTTGCAAGATCAAGAGAGTTTGTCTAAACTCATATGG	1140
Qy	1238	CAGGAAACTATGAAAGAGCCCTTGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCGAGG	1297
Db	1141	CAGGAAACTATGAAAGAGCCCTTGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCGAGG	1200
Qy	1298	TGTTGGATGTCACATCGATGATGGCATGCTAGATGGTCCAAAGTGCATGCAATGACAGATTTT	1357
Db	1201	TGTTGGATGTCACATCGATGATGGCATGCTAGATGGTCCAAAGTGCATGCAATGACAGATTTT	1260
Qy	1358	GCAACTTAAATGCTTCCGAGCCAGACATCGCAAGGTACCTTTGTGTCATCGACTCTCTCCA	1417
Db	1261	GCAACTTAAATGCTTCCGAGCCAGACATCGCAAGGTACCTTTGTGTCATCGACTCTCTCCA	1320
Qy	1418	ATTTTGTGCTGATTTGAAGCTGGGTTAAAGTGTGCCAAGGGAAAGTGCATTTGTCAATAGCA	1477
Db	1321	ATTTTGTGCTGATTTGAAGCTGGGTTAAAGTGTGCCAAGGGAAAGTGCATTTGTCAATAGCA	1380
Qy	1478	TTAGTCTGAAAGAGGAGAGGACGACTTTCTTGAGAGAGGCCAGGAGATTTAAAAGTATG	1537
Db	1381	TTAGTCTGAAAGAGGAGAGGACGACTTTCTTGAGAGAGGCCAGGAGATTTAAAAGTATG	1440
Qy	1538	GAGTCTGATGTTGCTCATGGCTTTTGTGTAAGAGGACAGGCAACAGAAACAGACACAA	1597
Db	1441	GAGTCTGATGTTGCTCATGGCTTTTGTGTAAGAGGACAGGCAACAGAAACAGACACAA	1500
Qy	1598	AAATCAGAGTGTGCACCGGGCTTACCATCTGCTGTGTGAAAAAATCTGGGCTTTAACTCAA	1657
Db	1501	AAATCAGAGTGTGCACCGGGCTTACCATCTGCTGTGTGAAAAAATCTGGGCTTTAACTCAA	1560
Qy	1658	ATGACATTTATTTTGAACCTTAATATCTAAACATTTGGGACTGGAATGGAGGAAACAACT	1717
Db	1561	ATGACATTTATTTTGAACCTTAATATCTAAACATTTGGGACTGGAATGGAGGAAACAACT	1620
Qy	1718	TGATGCAATTAATTTTATCCATGCAACAAAGTCAATTAAGAAACATTAACCTGGAGCCA	1777
Db	1621	TGATGCAATTAATTTTATCCATGCAACAAAGTCAATTAAGAAACATTAACCTGGAGCCA	1680
Qy	1778	GAATAAGTGGAGGTCTTTCCAACTTGTCTCTCTTCCAGGAAATGGAAAGCAATTCGAG	1837
Db	1681	GAATAAGTGGAGGTCTTTCCAACTTGTCTCTCTTCCAGGAAATGGAAAGCAATTCGAG	1740
Qy	1838	AAGCAATGCAATGGGTTTCTCTTTAATCCATGCAATCAAGTCTGGCATGGAGATAG	1897
Db	1741	AAGCAATGCAATGGGTTTCTCTTTAATCCATGCAATCAAGTCTGGCATGGAGATAG	1800
Qy	1898	TGAATGCTGGAAACCTCTCTGCTGTATGATATCCATAGGACTTCTGAGCTCTGTG	1957
Db	1801	TGAATGCTGGAAACCTCTCTGCTGTATGATATCCATAGGAACTTCTGAGCTCTGTG	1860
Qy	1958	AAGATCTCATCTGGAATAAAGACCTCAGGCACTGAGAAAGCTTTACGTTTATGCCAGA	2017
Db	1861	AAGATCTCATCTGGAATAAAGACCTCAGGCACTGAGAAAGCTTTACGTTTATGCCAGA	1920
Qy	2018	CTCAAGGCACAGGAGGAGGAAAGTCAATCAGACTGATGATGGAGAAATGGCCCTGTG	2077
Db	1921	CTCAAGGCACAGGAGGAGGAAAGTCAATCAGACTGATGATGGAGAAATGGCCCTGTG	1980
Qy	2078	AAGAACGCTTGAATATGCCCTTGTGAAAGGCAATGAAAGCAATTTATTTAGGATACTG	2137
Db	1981	AAGAACGCTTGAATATGCCCTTGTGAAAGGCAATGAAAGCAATTTATTTAGGATACTG	2040
Qy	2138	AGGAAGCCAGCTTAAACCAAAAAAATATCCCGACCTCTCAATATTAATTAAGAGACCCC	2197
Db	2041	AGGAAGCCAGCTTAAACCAAAAAAATATCCCGACCTCTCAATATTAATTAAGAGACCCC	2100
Qy	2198	TGATGAATGGAATGAAATTTGTTGGTGTCTTTTGGAGCTGGAAAAATGTTTCTACTC	2257
Db	2101	TGATGAATGGAATGAAATTTGTTGGTGTCTTTTGGAGCTGGAAAAATGTTTCTACTC	2160



```
Qy 2258 AGGTTATTAAGTACGCGCGGTATGAAAGAGGCTGTTGGCCACCTTATCCCTTTCATGG 2317
Dy |||||
2261 AGGTTATANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2220
Qy 2318 AAAAAGAAAGAGAAGAACAGAGTGTCTTAAACGGCACAGTAGAAGAGAGACCTTACC 2377
Dy |||||
2221 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2280
Qy 2378 AGGACCATCGTCTGGCCACTGTTAAAGCGCAGCTGTCAGACATAGGCAAGAACATAG 2437
Dy |||||
2281 AGGACCATCGTCTGGCCACTGTTAAAGCGCAGCTGTCAGACATAGGCAAGAACATAG 2340
Qy 2438 TTGAGTAGTCTTGGCTGCAATTAATTTCCGAGTTATTTAGTAGTCAATGACTCCAT 2497
Dy |||||
2341 TTGAGTAGTCTTGGCTGCAATTAATTTCCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2400
Qy 2498 GTGATAGACTACTGAAAGCTGCTCTTGACACAAAGCAGATATAATTGGCTGTCAGGAC 2557
Dy |||||
2401 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2460
Qy 2558 TCATCACTCTTCCCTGGATGAAATGATTTTGTGGCAAGGAAATGGAGATTAGCTA 2617
Dy |||||
2461 TCATCACTCTTCCCTGGATGAAATGATTTTGTGGCAAGGAAATGGAGATTAGCTA 2520
Qy 2618 TAAGGATTCATTTGTTGATTTGGAGAGCAACCACTTCAAAAACCCACACAGAGTTAAA 2677
Dy |||||
2521 TAAGGATTCATTTGTTGATTTGGAGAGCAACCACTTCAAAAACCCACACAGAGTTAAA 2580
Qy 2678 TAGTCTCGAGATACAGTGCACCTGTAATCCATGTCCTGGAGCGCTCCAGAGTGTGGT 2737
Dy |||||
2581 TAGTCTCGAGATACAGTGCACCTGTAATCCATGTCCTGGAGCGCTCCAGAGTGTGGT 2640
Qy 2738 TGTCTTCCAGCTGTAGATGAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAG 2797
Dy |||||
2641 TGTCTTCTCAGCTGTAGATGAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAG 2700
Qy 2798 AATATGAAGATATTAGACAGGACCATATGATGTTCTCAAGGAGAGGAGTACTTACCT 2857
Dy |||||
2701 AGTATGAAGATATTAGACAGGACCATATGATGTTCTCAAGGAGAGGAGTACTTACCT 2760
Qy 2858 TAAGTCAAGCAGCAAAAAGTGGTTCCAAATGGATGGCTGTCTGTAACCTCACCAGTGA 2917
Dy |||||
2761 TAAGTCAAGCAGCAAAAAGTGGTTCCAAATGGATGGCTGTCTGTAACCTCACCAGTGA 2820
Qy 2918 AGCCACGTTTATTGGGACCCAGTCTTTGAAGACTATGACCTCAGAGCTGTGGACT 2977
Dy |||||
2821 AGCCACGTTTATTGGGACCCAGTCTTTGAAGACTATGACCTCAGAGCTGTGGACT 2880
Qy 2978 ACATTGACTGGAAGCTTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAG 3037
Dy |||||
2881 ACATTGACTGGAAGCTTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAG 2940
Qy 3038 GCTTCCCAAGATATTAAACACAAACAGTAGTGGAGAGGAGGAGGAGTCTACCATG 3097
Dy |||||
2941 GCTTCCCAAGATATTAAACACAAACAGTAGTGGAGAGGAGGAGGAGTCTACCATG 3000
Qy 3098 ATGCCCAATATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCGGGGTGTGG 3157
Dy |||||
3001 ATGCCCAATATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCGGGGTGTGG 3060
Qy 3158 TTGGGTTCTGGCCAGCACAGAGTATCCAAAGACGACATTACCTGTAGCGAGAGCTGTG 3217
Dy |||||
3061 TTGGGTTCTGGCCAGCACAGAGTATCCAAAGACGACATTACCTGTAGCGAGAGCTGTG 3120
Qy 3218 TGCCCCAGGCTGCAGAGCCCATAGCCATTTCTATGGGTTAAGGCAACAGGCTGAGAAGG 3277
Dy |||||
3121 TGCCCCAGGCTGCNNNGCCCATAGCCATTTCTATGGGTTAAGGCAACANNNNNNNNNNNN 3180
Qy 3278 ACTCTGCCAGCAGGAGCCATCTACTGCTCTCAGACTTCATCGCTCCCTTGCATTCTG 3337
Dy |||||
3181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3240
```

```
Qy 3338 GCATCCGCTGACTACCTGGGCGCTGTTTCCGCTGCTCTTTGGGTAGAGAGCTGAGCA 3397
Dy NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3300
Qy 3398 AGGCTTATGAGGATGATGGTGACGACTACAGCAGCATCATGGTCAAGCGCTGGGGGACC 3457
Dy |||||
3301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3360
Qy 3458 GGCTGGCAGAGGCTTTTGCAAGAGAGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCT 3517
Dy |||||
3361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3420
Qy 3518 ACTGTGGCAGTACAGCAGCTGCGTGGCAGACCTGCGAAGGTTGCGGTACAGGGCATCC 3577
Dy |||||
3421 ACTGTGGCAGTACAGCAGCTGCGTGGCAGACCTGCGAAGGTTGCGGTACAGGGCATCC 3480
Qy 3578 GCCCGGCTCTGCTACCCCGCAGCAGCCGACACACCCGAGAGCTCACCATGTGGAGAC 3637
Dy |||||
3481 GCCCGGCTCTGCTACCCCGCAGCAGCCGACACACCCGAGAGCTCACCATGTGGAGAC 3540
Qy 3638 TCCGAGACATCGAGCAGTCTTACAGGCAATAGGTTAAACAGAACTATTAGCAATGGCACCTG 3697
Dy |||||
3541 TCCGAGACATCGAGCAGTCTTACAGGCAATAGGTTAAACAGAACTATTAGCAATGGCACCTG 3600
Qy 3698 CTTTCAGAGTCTCAGGCTCTACTTCTCAATTTGAAGTCAAAATATTTCTGTGGGGA 3757
Dy |||||
3601 CTTTCAGAGTCTCAGGCTCTACTTCTCAATTTGAAGTCAAAATATTTCTGTGGGGA 3660
Qy 3758 AGATTTCCAGGATCAG 3774
Dy |||||
3661 AGATTTCCAGGATCAG 3677

RESULT 3
CR859882
LOCUS Pongo pygmaeus mRNA; cDNA DKFp469P1116 (from clone DKFp469P1116).
DEFINITION Pongo pygmaeus mRNA; cDNA DKFp469P1116 (from clone DKFp469P1116).
ACCESSION CR859882
VERSION CR859882.1 GI:55730633
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pongo.
REFERENCE 1 (bases 1 to 4888)
AUTHORS Otenwaelder,B.; Obermaier,B.; Deutschenbaeur,S.; Schaipp,A.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFp469P1116) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFp469P1116
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..4888
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFp469P1116"
/tissue_type="kidney"
/clone_lib="469 (synonym: pkidl). Vector pSport1_Sfi; host
DH10B; sites SfilA + SfilB"
/dev_stage="adult"
```

```

/notes~similar to 5-methyltetrahydrofolate--homocysteine
methyltransferase (Homo sapiens)"
1. 4888.
/genes~DKFZp469P1116"
595..2754
/genes~DKFZp469P1116"
/codon_start=1
/product~"hypothetical protein"
/protein_id~"CAH92038.1"
/db_xref~"GI:55730634"
/translation~"MEHNLYAINFIHATKVIKETLPGARISGNSLNGSFRGMEALR
REAMQGVFIYHAIKSGDMGIVNAGNLNPEVYDIIKELLQICLDIWNKQPEATEKLLR
YAQTQGGKMGKIVDTWRNAGSREISYALVKGXIHVEDTEARLAEKQEPYRLIN
IISGLPMGKMKIVGLDFGAKMFLPOVTKSARVMKKAQVHIDPEMEKAREETVRLNG
VERSEDPQSTGLIVLAVTKGVDPVHDIGKNIIVGVVGLNPNRFDVLGWTVPDCKILKAALDH
KADIIQGLTGLIIPSELDITFVAKEMNIAKIPILIGATSTHTAVKTAIPYASAPV
IHLVDSKSVVCSQLLDENLKDPEIIMEEYEDIQDHVESLKERRYILPLSQARKS
FGQMDLMSBPVKVPTQITQFVEDYIDQKVLQVDMKVPFVQWQKGLTPNKGFPKI
FNQKTMVSGEAKYVITDQAQMNLNTLSKQKLQARGVQWFAQSTQDDIHLYABGAVPQ
AAEDPTATVYGLRQAKGDSASTPESTIAPLHGIRDYLGLPFAVACFGVEELSK
AYEDPDGDTYSSIMVAKGLDARAEAFBELHVERVLRNWAYCVSGEQDLADVRLRYKQ
IRPAGVPQSDPQETKLTWMLRADIEQSTGRTLTETSLMA PASAGSLVPSNLKSTYF
AVCKTSKQNDVYALPKNTSVAREVKMTGPTIGVYDNTM"

```

## ORIGIN

Query Match	58.2%	Score 2282.6	DB 4	Length 4888	
Best Local Similarity	92.6%	Pred. No. 0			
Matches 2485	Conservative	0	Mismatches 39	Indels 159	Gaps 2
Qy	1391	AGGTACCTTTGTGCATCGACTCTCCCAATTTTGTCTGTGATTGAAGCTGGGTTAAAGTGCT	1450		
Db	130	AGGTACCTTTGTGCATCGACTCTCCCAATTTTGTCTGTGATTGAAGCTGGGTTAAAGTGCT	189		
Qy	1451	GCCAAAGGGAAGTGCAATTGTCAATAGCAATTAGTCTGAAGGAAGGAGAGGACGACTTTCTGG	1510		
Db	190	GCCAAAGGGAAGTGCAATTGTCAATAGCAATTAGTCTGAAGGAAGGAGAGGACGACTTTCTGG	249		
Qy	1511	AGAAGGCCACGAGAGATTTAAAGATATGGAGCTGCTATGGTGGTCAATGGCTTTTGTATGAAG	1570		
Db	250	AGAAGGCCGGAAGATTTAAAGATATGGAGCTGCTATGGTGGTCAATGGCTTTTGTATGAAG	309		
Qy	1571	AAGGACA-----	1577		
Db	310	AAGGACAGATACATTTTCTTTCAATTGTGAAAAAATTCAGATCATCCAGAGAATGCAAAAG	369		
Qy	1578	-----	1577		
Db	370	AGTGATAGTTTGTGATCAATTAACATATATTGAACCTTAAATTCACACCAGCAAGTACACAA	429		
Qy	1578	-----GGCAACAGAAAAACAGACACA	1596		
Db	430	CACACATTAATACAGTTCTTCTCCAAGCATATAAATATATGGCAACAGAAAAACAGACACA	489		
Qy	1597	AAAATCAGAGTGGCAACCGGGCCCTACATCTGCTGTGTGAAAAAATCTGGGCTTTAATCCA	1656		
Db	490	AAAATCAGAGTGTGCACCGGGCCCTACATCTGCTGTGTGAAAAAATCTGGGCTTTAATCCA	549		
Qy	1657	AATGACATTAATTTTGTGACCTTAATATCTTAACCAATGGGACTGGAAATGGAGGAACACAAC	1716		
Db	550	AATGACATCATTTTGTGACCTTAATATCTTAACCAATGGGACTGGAAATGGAGGAACACAAC	609		
Qy	1717	TTGTATGCCATTAATTTTATCCATGCAACAAAAGTCAATTAAGAAAACATTTACTCTGGAGCC	1776		
Db	610	TTGTATGCCATTAATTTTATCCATGCAACAAAAGTCAATTAAGAAAACATTTACTCTGGAGCC	669		
Qy	1777	AGAAATAGTGGAGTCTTTTCCAACTGTCTTCTCTCCGAGAAATGGAGGCATTCGA	1836		
Db	670	AGAAATAGTGGAGTCTTTTCCAACTGTCTTCTCTCCGAGAAATGGAGGCATTCGA	729		
Qy	1837	GAAGCAATGCAATGGGGTTTTCTTTTACCATGCAATCAAGTCTGGCATGGACATGGAGATA	1896		
Db	730	GAAGCAATGCAATGGGGTTTTCTTTTACCATGCAATCAAGTCTGGCATGGACATGGAGATA	789		

1897	Qy	GTGAATGCTGGAAA	CCTCCCTGTGTATGATGATATCCA	TAAAGAACTTCTG	CAGCTCTCT	1955
790	Db	GTGAATGCTGGAAA	CCCTCCCTGTGTATGATGATATCCA	TAAAGAACTTCTG	CAGCTCTCT	849
1957	Qy	GAAGATCTCATCTGGAAT	TAAAGACCTCTGAGGCCACTTGAGAAGCTCTT	TAGCTTATGCC	CAG	2016
850	Db	GAAGATCTCATCTGGAAT	TAAAGACCTCTGAGGCCACTTGAGAAGCTCTT	TAGCTTATGCC	CAG	909
2017	Qy	ACTCAAGGCACAGGAGG	GAAGAAAGTCATTCAGACTGATGAGTGAGG	AAATGGCCCTGTC		2076
910	Db	ACTCAAGGCACAGGAGG	GAAGAAAGTCATTCAGACTGATGAGTGAGG	AAATGGCCCTGTC		969
2077	Qy	GAAGAACGCTTGAGTAT	GCCCTTGTGAAGGGGATTTGAAAAACATATTA	TATGAGGATACT		2136
970	Db	GAAGAACGCTTGAGTAT	GCCCTTGTGAAGGGGATTTGAAAAACATATTA	TATGAGGATACT		1029
2137	Qy	GAGGAAGCCAGGTTAA	ACCAAAAAAATATCCCGACCTCTCAATATAAT	TGGAAGACCC		2196
1030	Db	GAGGAAGCCAGGTTAA	ACCAAAAAAATATCCCGACCTCTCAATATAAT	TGGAAGACCC		1089
2197	Qy	CTGATGAATGGAATG	AAAAATTTGTTGGTGATCTTTTTTGGAGCTTGG	AAAAATGTTTCTACCT		2256
1090	Db	CTGATGAATGGAATG	AAAAATTTGTTGGTGATCTTTTTTGGAGCTTGG	AAAAATGTTTCTACCT		1149
2257	Qy	CAGGTTATAAGTC	AGCCCGGGTTATGAAGAAGGCTGTTGGCCACCTT	ATCCCTTCATG		2316
1150	Db	CAGGTTATAAGTC	AGCCCGGGTTATGAAGAAGGCTGTTGGCCACCTT	ATCCCTTCATG		1209
2317	Qy	GAAGAAGAGGAAGAA	ACCAGAGTCGTTAACGGCACAGTAGTAGAAGAG	AGACCCCTTAC		2376
1210	Db	GAAGAAGAGGAAGAA	ACCAGAGTCGTTAACGGCACAGTAGTAGAAGAG	AGACCCCTTAC		1269
2377	Qy	CAGGGCACCATCTG	CTGTGGCCACTGTTTAAAGGCGACGTGCACACAT	ATAGGCAAGAACATA		2436
1270	Db	CAGGGCACCATCTG	CTGTGGCCACTGTTTAAAGGCGACGTGCACACAT	ATAGGCAAGAACATA		1329
2437	Qy	GTTGGAGTAGTCTT	GGCTGCAATAATTTCCGAGTTATTGATTTAGGAGT	CACTGACTCCA		2496
1330	Db	GTTGGAGTAGTCTT	GGCTGCAATAATTTCCGAGTTATTGATTTAGGAGT	CACTGACTCCA		1389
2497	Qy	TGTGATAAGATAC	TGGAAGCTGCTTGACCA	CAAAAGCAGATATAATTCG	CGCTGTCAGGA	2556
1390	Db	TGTGATAAGATAC	TGGAAGCTGCTTGACCA	CAAAAGCAGATATAATTCG	CGCTGTCAGGA	1449
2557	Qy	CTCATCACTCTT	CCCTGGATGAAATGATTTTTTGTGGCCAAAGG	AAATGGAGAGATTAGCT		2616
1450	Db	CTCATCACTCTT	CCCTGGATGAAATGATTTTTTGTGGCCAAAGG	AAATGGAGAGATTAGCT		1509
2617	Qy	ATAAGGATTCAT	TGTTGATGGAGGAGCAACCACTTCAAAAA	CCCCACACAGCAGTTAAA		2676
1510	Db	ATAAGGATTCAT	TGTTGATGGAGGAGCAACCACTTCAAAAA	CCCCACACAGCAGTTAAA		1569
2677	Qy	ATAGCTCCGAGATAC	AGTGACCTGTAAATCCATGCTCTGGAACGCGT	TCCAAGAGTGTGGTG		2736
1570	Db	ATAGCTCCGAGATAC	AGTGACCTGTAAATCCATGCTCTGGAACGCGT	TCCAAGAGTGTGGTG		1629
2737	Qy	GTGTGTTCCAGCTG	TTAGATGAAATCTTAAAGGATGAATATCTT	TTCAGGAAAAATCATGGAA		2796
1630	Db	GTGTGTTCTCAGCT	TTAGATGAAATCTTAAAGGATGAATATCTT	TTCAGGAAAAATCATGGAA		1689
2797	Qy	GAATATGAAGATAT	TAGACAGGACCAATATGAGTCTCTCAAGGAG	AGGAGATACTTACCC		2856
1690	Db	GAATATGAAGATAT	TAGACAGGACCAATATGAGTCTCTCAAGGAG	AGGAGATACTTACCC		1749
2857	Qy	TTAAGTCAAGCCAG	AAAAAGTGGTTTCCAAATGGAATGGTGTCT	GTGAACTTCAACCCAGTG		2916
1750	Db	TTAAGTCAAGCCAG	AAAAAGTGGTTTCCAAATGGAATGGTGTCT	GTGAACTTCAACCCAGTG		1809
2917	Qy	AAGCCACAGTTTAT	TGGGACCCAGGCTCTTTTGAAGACTAT	TGACCTGCAAGAGCTGGTGGAC		2976
1810	Db	AAGCCACAGTTTAT	TGGGACCCAGGCTCTTTTGAAGACTAT	TGACCTGCAAGAGCTGGTGGAC		1869
2977	Qy	TACATTGACTGGA	AGCCCTTTCTTTTGTATGCTCGGAGCT	CCGGGGCAAGTACCCGAATCGA		3036

Db	1870	TACATTGACCTGGAAGCCCTTCTTTGATGTCTGGCAGCTCCGGGCAAGTACCCGAAATCGA	1929
Qy	3037	GGCTTCCCAAGATATTTAAACGACAAAACAGTACGTAGCTGGAGAGGCCAGGAAGGTTCTACGAT	3096
Db	1930	GGCTTCCCAAGATATTTAAACGACAAAACAGTACGTAGCTGGAGAGGCCAAAAGGTTCTACGAT	1989
Qy	3097	GATGCCCAACAATATGCTGGAACACACTGATTTAGTCAAAAGAAAATCTCCGGGCCCGGGGTGTG	3156
Db	1990	GATGCCCAAAAATATGCTGAACACACTGATTTAGTCAAAAGAAAATCTCCAAGCCCGGGGTGTG	2049
Qy	3157	GTTCGGGTTCTGGCCAGACACAGATATCCAGACGACATTCACCTGTACGCAGAGGCTGCT	3216
Db	2050	GTTCGGGTTCTGGCCAGACACAGATATCCAGATGACATTCACCTGTACGCAGAGGGTGCT	2109
Qy	3217	GTGCCCCAGGCTGCAGAGCCCATAGCCACTTTCTATGGGTTAAAGCCAACAGGCTGACAAG	3276
Db	2110	GTGCCCCAGGCTGCAGAGCCCATAGCCACTTTCTATGGGTTAAAGCCAACAGGCTGAGAG	2169
Qy	3277	GACTCTGCCAGCACGGAGCCCATACTACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCT	3336
Db	2170	GACTCTGCCAGCACGGAGCCCATACTACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCT	2229
Qy	3337	GGCATCCGTGACTACCTGGGCCCTGTTTGGCCGTTGCCCTGCTTTGGGTTAGAAGACTGAGC	3396
Db	2230	GGCATCCGTGACTACCTGGGCCCTGTTTGGCCGTTGCCCTGCTTTGGGTTAGAAGACTGAGC	2289
Qy	3397	AAGCCCTATGAGGATGATGGTGACGACTACAGCAGCATCATGGTCAAGGCGCTGGGGGAC	3456
Db	2290	AAGCCCTATGAGGATGATGGTGACGACTACAGCAGCATCATGGTCAAGGCGCTGGGGGAC	2349
Qy	3457	CGGCTGGCAGAGGCTTTTGCAGAAAGACTTCCATGAAGAAGTTCGCCCGAGAACTGTGTGGCC	3516
Db	2350	CGGCTGGCAGAGGCTTTTGCAGAAAGACTTCCATGAAGAAGTTCGCCCGAGAACTGTGTGGCC	2409
Qy	3517	TACTGTGGCAGTGACAGCTGGAGCTGCGAGACCTGCGAAGGTTGCGGTHACAAGGCGATC	3576
Db	2410	TACTGTGGCAGTGACAGCTGGAGCTGCGAGACCTGCGAAGGTTGCGGTHACAAGGCGATC	2469
Qy	3577	CGCCCGGCTCTCTGGCTACCCCGAGCCAGCCGAGACCAACGAGAGGCTCACCATGTGGAGA	3636
Db	2470	CGCCCGGCTCTCTGGCTACCCCGAGCCAGCCGAGACCAACGAGAGGCTCACCATGTGGAGA	2529
Qy	3637	CTGCAGACATCGAGCAGTCTACAGGCAATTAGGTTAAACAGAACTCATTAGCAATGGCACCT	3696
Db	2530	CTGCAGACATCGAGCAGTCTACAGGCAATTAGGTTAAACAGAACTCATTAGCAATGGCACCT	2589
Qy	3697	GCTTCAGAGTCTCAGGCGCTCTACTTCTCCAATTTGAAGTCCAATAATTTGCTGTGGGG	3756
Db	2590	GCTTCAGAGTCTCAGGCGCTCTACTTCTCCAATCTGAAGTCCAATATTTGCTGTGGGG	2649
Qy	3757	AAGATTTCCAAAGGATCAGGTTGAGGATATGCAATGAGGAAGAACATATCTGTGGCTGAG	3816
Db	2650	AAGATTTCCAAAGGATCAGGTTGAGGATATGCAATGAGGAAGAACATATCTGTGGCTGAG	2709
Qy	3817	GTTCAGAAATGGCTTTGGACCCATTTGGGATATGATACAGACTAACTTTTTTTTTTTTTT	3876
Db	2710	GTTCAGAAATGGCTTTGGACCCATTTGGGATATGATACAGACTTAAC-----TTTTTTTTT	2764
Qy	3877	TTGCCTTTTTTATCTTGATGATCTCAAGGAAATACAACTTAG	3919
Db	2765	TGCCTTTTTTGTCTTGATGATCTCAAGGAAATACAACTTAG	2807
RESULT 4			
AK037599			
LOCUS			
DEFINITION			
Mus musculus 16 days neonate thymus cDNA, linear HTC 03-APR-2004			
enriched library, clone: A130029B16			
product: 5-methyltetrahydrofolate-homocysteine methyltransferase,			
full insert sequence.			
ACCESSION			
VERSION			
AK037599.1 GI:26085945			



Qy 1989 CACTGAGAGCTCTTACGTATGCCCAGACTCAAGGCACAGGAGGAGGAGTCA 2044  
 Db 2082 TACTGAGAACTCTTGCATATGCCAGACTCATGGCAGAGGAGGAGTAA 2137

RESULT 5  
 AK085987 LOCUS  
 DEFINITION Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830038K18 product:hypothetical Vitamin B12 dependent methionine synthase activation domain containing protein, full insert sequence.

ACCESSION AK085987  
 VERSION AK085987.1 GI:26351896  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 10349636

2  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 11042159

3  
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitasuna,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 11076861

4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)

5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

6 (bases 1 to 1788)  
 Adachi,J., Aizawa,K., Akimura,T., Atakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saiki,R., Saich,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaihi,S., Tanaka,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (15-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,  
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/.

FEATURES  
 Location/Qualifiers  
 source  
 1..1788  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:D830038K18"  
 /db\_xref="taxon:10090"  
 /clone="D830038K18"  
 /tissue\_type="heart"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="16 days neonate"  
 <1..1318  
 /note="unnamed protein product; hypothetical Vitamin B12 dependent methionine synthase activation domain containing protein (InterPro|IPR004223, evidence: InterPro)  
 putative"  
 /codon\_start=2  
 /protein\_id="BAC39585.1"  
 /db\_xref="GI:26351897"  
 /translation="VRLPMNLGSLIHSFSGISGWEKFLVYKRSVKSACKSPVR  
 HCSDDVNSGSELSALTDVLSFWTLTRTHTAVKIAPRYSAPIVHLVDASKSVVYRK  
 VPLSQARKGHFIDMLSEPHVPTFGTGFEDYLNKLVDYIDKPFDFVWQDRGK  
 VYNGRPKIFNDKAVGEARKVNDQNMNLILISOKKQARGVGVFWPAQSVQDDIH  
 LYAEGVVQAAPFIATFYGLRQQAEDSSDTPVHCLSDFIAPLHSGVCYVLGFAVA  
 CFVGEELSKTYEDDGDYSSIMYKAGDLRLAEAPAEHLHERVRELWAYRSQGLGVP  
 DLRLREGIRPAPGPSQDTEKULMRLASTEQTGRLTRESLAWAPASAVSGLY  
 FSNVKAIFYAVGKISKDQTEDYALRNKMPVAEVEKMLGPLIGYDTD"  
 1767..1772  
 /note="putative"  
 1788  
 /note="putative"

polya\_signal  
 polya\_site  
 ORIGIN  
 Query Match 20.9%; Score 819.2; DB 4; Length 1788;  
 Best Local Similarity 87.5%; Pred. No. 2.4e-217;  
 Matches 896; Conservative 0; Mismatches 128; Indels 0; Gaps 0;  
 Qy 2838 GGAGAGGAGATCTTACCTTAAGTCAAGCCAGAGAAAGTGGTTTCCAAATGGATTGGCT 2897  
 Db 295 GGAGAGAAATACGTACCCCTTAAGTCAAGCCAGAGAAACACGGTTTCCACATTGATTGGCT 354  
 Qy 2898 GTCTGAACCTCACCCAGTGAAGCCAGCTTATTGGACCAGCCAGTCTTTGAAGACTATGA 2957  
 Db 355 CTCTGAACCTCATCCAGTGAAGCCAGCTTATTGGACCAGCCAGTCTTTGAGGACTACAA 414  
 Qy 2958 CTTGCAAGCTGGTGGACTACATTGACTGGAAGCCCTTCTTGTGATCTCTGGCAGCTCCG 3017  
 Db 415 CTTGCAAGCTGGTGGACTACATTGACTGGAAGCCCTTCTTGTGATCTCTGGCAGCTCCG 474  
 Qy 3018 GGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTAAACGACAAAAACAGTAGGTGGAGA 3077  
 Db 475 GGGCAAGTACCCGAAATCGAGGATTTCCCAAGATATTTAAATGACAAACAGTAGGTGAAGA 534  
 Qy 3078 GGGCAGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTGATTAGTCAAAGAA 3137  
 Db 535 AGCCAGGAAGGTATACAATGATGCTCAGAAATATGCTGGAACATATCTGATTAGTCAAAGAA 594  
 Qy 3138 ACTCCGGGCCCCGGGTGTGGTCTGCGCCAGCAGAGTATCCAAAGCAGCATTTCA 3197  
 Db 595 ACTGAGGCCAGGGGTGTGGTGTGGATTCTGGCAGCAGAGTGTCCAGGATGATCACTCA 654  
 Qy 3198 CTTGTACGAGAGGCTCTGTGCCCCCAGGCTGCAGAGCCCATAGCCACTTTCTATGGGTT 3257



655 CCTGTATGACAGGGGGTGTGCCCGCGCTGAGCCATAGCCACTTCTATGGACT 714  
3358 AAGCAACAGCTGAGAGGACTCTGCGACAGCGGACCATACTACTGCTCTCAGACTT 3317  
715 GAGCAGCAGGCTGAGAGGACTCTTCTAGTACAGACCCCTACCAGCTCTCAGACTT 774  
3318 CATCGCTCCCTTGCATTCTGGCATCCGCTGACTTACTGCGGCTGTTTGGCGTTCGCTGCTT 3377  
775 CATTCCTCTTGCATTCTGGTGTCTGTGACTACTTGGGTCTGTTTGGTGTCTGCTGCTT 834  
3378 TGGGGTGAAGAGCTGAGAGGCTATGAGGATGATGGTGCAGCACTACAGCAGCATCAT 3437  
835 TGGGGTGAAGAGCTGAGTGAAGACTATGAGGATGATGGGCTGATGACTACAGCAGCATCAT 894  
3438 GGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTCAGAGAGCTCCATGAAGAGT 3497  
895 GGTGAAGCGCTGGGCGAGCGCTGGCAGAGGCTTTCAGAGAGCTCCATGAGAGT 954  
3498 TCGCGGAGACTGTGGGCTTACTGTGGCAGTGAGCAGCTGAGCTTCGAGACCTGCCAAG 3557  
955 TCGCGGAGACTGTGGGCTTACTGTAGAGTGAGCAGTGGGCTTCCAGACTTGGCAG 1014  
3558 GTTCGGGTACAAGGGCATCCCGCGCTCTGCTGCTACCCAGCCAGCCGACACCGA 3617  
1015 ACTCGGTATGAGGGCATCCGCGCAGCTCTGCTTACCCAGCCAGCCGCTGACCATCTGA 1074  
3618 GAAGCTACCATGTGGAGACTCGCAGACATCGAGGACTCTACAGGCTTACAGCA 3677  
1075 GAAGCTTACCATGTGGAGACTGGCAGCATCGAGCAGCCACAGGCTGACGACAGA 1134  
3678 ATCATAGCAATGGCAGCTGTGAGCAGTCTCAGGCTCTACTTCTCCAAATTTGAAGTC 3737  
1135 ATCTTGGCAATGGACCCGCTTCAGCAGTGTGAGGCTGTACTTCTCCAACTGAAGGC 1194  
3738 CAATATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTGAGGATTAATGATGAGGAA 3787  
1195 CAATATATTTGCTGTGGGGAAGATTTCCAAAGATCAGACTGAGGATTAATGATGAGGAA 1254  
3798 GAACATATCTGTGCTGAGGTGAGAAATGCTGGACCCATTTTGGATATGATACAGA 3857  
1255 GAACATGCGAGTGGCGAGGTGGAGAAATGGCTTGGTCCCATTTCTGGGCTATGACACAGA 1314  
3858 CTAA 3861  
1315 CTGA 1318

RESULT 6  
BM450887  
LOCUS BM450887\_6387311 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5490921  
DEFINITION 5', mRNA sequence.  
ACCESSION BM450887  
VERSION BM450887.1 GI:18499927  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
1 (bases 1 to 863)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
Plate: LLAM12110 row: j column: 10  
High quality sequence stop: 748.  
Location/Qualifiers  
source 1. 863  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5490921"  
/tissue\_type="retinoblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_67"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: Not1;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Library constructed by Life  
Technologies."  
ORIGIN  
Query Match 20.8%; Score 814.8; DB 3; Length 863;  
Best Local Similarity 98.9%; Pred. No. 3.2e-216;  
Matches 841; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
Qy 283 CAGCTGATGTCAATTAACCAATCCATAAGAAATACCTGCTGGCTGGGGCAGATATCATTT 342  
Db 11 CGGCTGATGTCAATTAACCAATCCATAAGAAATACCTGCTGGCTGGGGCAGATATCATTT 70  
Qy 343 GAAACAAATCTTTAGCAGCACTAGTATTGCCCAAGCTGACTATGCGCTTGAACACTTG 402  
Db 71 GAAACAAATCTTTAGCAGCACTAGTATTGCCCAAGCTGACTATGCGCTTGAACACTTG 130  
Qy 403 GCCTACCGGATGAACATGTGCTCTGCGAGGATGGCCAGAAAAGCTGCGAGGAGGTAACT 462  
Db 131 GCCTACCGGATGAACATGTGCTCTGCGAGGATGGCCAGAAAAGCTGCGAGGAGGTAACT 190  
Qy 463 CTCAGACAGAAATTAAGAGGTTTGTGGCAGGGCTCTGGGTCCGACTAATAAGACACTC 522  
Db 191 CTCAGACAGAAATTAAGAGGTTTGTGGCAGGGCTCTGGGTCCGACTAATAAGACACTC 250  
Qy 523 TCTGTGTCCTCATCTGTGGAAAGCCGATTTATAGGAACATCAATTTGATGAGCTTGT 582  
Db 251 TCTGTGTCCTCATCTGTGGAAAGCCGATTTATAGGAACATCAATTTGATGAGCTTGT 310  
Qy 583 GAACATACCAAGACAGGCGCAAGGACTTCTGATGCGCGGGTTGATATCTTACTCATTT 642  
Db 311 GAACATACCAAGACAGGCGCAAGGACTTCTGATGCGCGGGTTGATATCTTACTCATTT 370  
Qy 643 GAAACTATTTTGTATCTGCCAATGCCAAGCAGCTTGTGTGCACTCCAAAATCTTTTT 702  
Db 371 GAAACTATTTTGTATCTGCCAATGCCAAGCAGCTTGTGTGCACTCCAAAATCTTTTT 430  
Qy 703 GAGGAGAAATATGCTCCCGCCCTATCTTTATTTTCAGGAGCGATCGTTGATAAAGTGGG 762  
Db 431 GAGGAGAAATATGCTCCCGCCCTATCTTTATTTTCAGGAGCGATCGTTGATAAAGTGGG 490  
Qy 763 CGGACTCTTTCGGACAGACAGGAGGAGGATTTGTCTATCAGCGGTGTCTATGGAGAACCA 822  
Db 491 CGGACTCTTTCGGACAGACAGGAGGAGGATTTGTCTATCAGCGGTGTCTATGGAGAACCA 550  
Qy 823 CTCTGCATTTGATTAATTTGTCTTTGGTGGCAGCTGAGATGAGACCTTTTATTTGAAATA 882  
Db 551 CTCTGCATTTGATTAATTTGTCTTTGGTGGCAGCTGAGATGAGACCTTTTATTTGAAATA 610  
Qy 883 ATTGGAATTTGATCAACAGCCTATGCTCTCTGTTATCCCAATGAGGCTTCCCAACACC 942  
Db 611 ATTGGAATTTGATCAACAGCCTATGCTCTCTGTTATCCCAATGAGGCTTCCCAACACC 670  
Qy 943 TTTGCTGACTATGATGAAAAGCCCTTCTATGATGGCCAGCACTTAAAGGATTTTGTCTATG 1002  
Db 671 TTTGCTGACTATGATGAAAAGCCCTTCTATGATGGCCAGCACTTAAAGGATTTTGTCTATG 730  
Qy 1003 GATGGCTTGGTCAATATAGTTTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGAA 1062  
Db 731 GATGGCTTGGTCAATATAGTTTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGAA 790



```

Qy 1063 ATTCTGAAGCTGTGAATAATTTAAG-CCTAGAGTTCCACCTGCCA-CTGCTTTTGAAG 1120
Db 791 ATTGCTCTGAGCTGTGAATAATTTAAGCCCTAGAGTTCCACCTGCCACCTCTTTTGAAG 850
Qy 1121 GACATATGTT 1130
Db 851 GACATATGTT 860

RESULT 7
BQ062840
LOCUS BQ062840 1043 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6826659 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5924378
5', mRNA sequence.
ACCESSION BQ062840
VERSION BQ062840.1 GI:19890017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1043)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2094 row: 9 column: 03
High quality sequence start: 242
High quality sequence stop: 479.
Location/Qualifiers
1. 1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5924378"
/tissue type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

FEATURES
source
1. 1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5924378"
/tissue type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 20.8%; Score 814; DB 3; Length 1043;
Best Local Similarity 98.2%; Pred. No. 5.8e-216;
Matches 876; Conservative 0; Mismatches 10; Indels 6; Gaps 5;

Qy 2422 ATAGCAAGACATAGTTGGAGTAGTCCTT-GGCTGCAATAATTTCCGAGTTATTGATTT 2480
Db 56 ATAGCAAGACATAGTTGGAGTAGTCCTTGGCTGCAATAATTTCCGAGTTATTGATTT 115
Qy 2481 AGGAGTCATGACTCCATGCTGATAAGATACTGAAAGCTGCTCTTGACCAACAAGCAGATAT 2540
Db 116 AAGAGTCATGACTCCATGCTGATAAGATACTGAAAGCTGCTCTTGACCAACAAGCAGATAT 175
Qy 2541 AATTGGCTGTGAGAGTCATCATCTCTTCCCTGGATGAATGATTTTGTGTCGAAGGA 2600

```

```

Db 176 AATTGGCTGTGAGAGTCATCATCTCTTCCCTGGATGAATGATTTTGTGTCGAAGGA 235
Qy 2601 AATGGAGAGATTAGCTATAAGGATTCATTGTTGATTGGAGGACCAACCATCTCAAAAC 2660
Db 236 AATGGAGAGATTAGCTATAAGGATTCATTGTTGATTGGAGGACCAACCATCTCAAAAC 295
Qy 2661 CCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTGTATCCATGTCCTGGAGGC 2720
Db 296 CCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTGTATCCATGTCCTGGAGGC 355
Qy 2721 GTCCAAGAGTGTGCTGTGTTTCCACAGCTGTTAGATGAAAATCTAAAGGATGAATCTT 2780
Db 356 GTCCAAGAGTGTGCTGTGTTTCCACAGCTGTTAGATGAAAATCTAAAGGATGAATCTT 415
Qy 2781 TGAGGAAATCATGAAAGATATGAAGATATTAGACAGGACATATTATGAGTCTCTCAAG-G 2839
Db 416 TGAGGAAATCATGAAAGATATGAAGATATTAGACAGGACATATTATGAGTCTCTCAAGAG 475
Qy 2840 AGAGGAGATACTTACCCTTAAGTCAAGCCAGAAAAGTGTTCCTCAATGGAATGTCCTGT 2899
Db 476 AGAGGAGATACTTACCCTTAAGTCAAGCCAGAAAAGTGTTCCTCAATGGAATGTCCTGT 535
Qy 2900 CTGAACCTCTCAACCCAGTGAAGCCACGTTATTGGGACCCAGGCTCTTTGAAGACTATGACC 2959
Db 536 CTGAACCTCTCAACCCAGTGAAGCCACGTTATTGGGACCCAGGCTCTTTGAAGACTATGACC 595
Qy 2960 TGCAAGAGCTGTGGACTACATTGACTGGAAGCCCTTTCTTTGATGTCTGGCAGCTCCGGG 3019
Db 596 TGCAAGAGCTGTGGACTACATTGACTGGAAGCCCTTTCTTTGATGTCTGGCAGCTCCGGG 655
Qy 3020 GCAAGTACCCGATTCGAGGCTTCCCAAGATATTTAACGACAAAACAGTAGTGGAGAGG 3079
Db 656 GCAAGTACCCGATTCGAGGCTTCCCAAGATATTTAACGACAAAACAGTAGTGGAGAGG 715
Qy 3080 CCAGGAAGGCTCTACGATGATGCCACAAATATGCTGGAACACACTGATTAGTCAAAAGAAC 3139
Db 716 CCAGGAAGGCTCTACGATGATGCCACAAATATGCTGGAACACACTGATTAGTCAAAAGAAC 775
Qy 3140 TCCGGGGCCCGGGG-TGTGGTTGGGTTCTGCGCAGCAGAGTAGTATCCAGACGACATTAC 3198
Db 776 TCCGGGGCCCGGGGNTGTTGGTTGTTGTTCTGCGCAGCAGAGTAGTATCCAGACGACATTAC 835
Qy 3199 CTGTACCCAGAGGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3258
Db 836 CTGTACCCAGAGGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 895
Qy 3259 AGGCAACA--GGCTGGAAGGACTCTGCCAGCAGCG-AGCCATACTACTGCC 3307
Db 896 AGGCAACAAGGCTGGAGAGGACTCTGCCAGCAGCGNAACCATCTACTACTGCC 947

RESULT 8
BX390174
LOCUS BX390174 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0C010Y017 5-PRIME, mRNA sequence.
ACCESSION BX390174
VERSION BX390174.2 GI:46877980
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 871)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30463266.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 740.f

For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?8=CS0BAG006ZC05\\_CS00515\\_1&c=740.f](http://www.genoscope.cns.fr/cdna?8=CS0BAG006ZC05_CS00515_1&c=740.f).

#### FEATURES

Location/Qualifiers

1..871

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DC010Y017"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

```
Query Match      20.3%; Score 797.4; DB 5; Length 871;
Best Local Similarity 98.0%; Pred. No. 2.5e-211;
Matches 837; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 669 CAAGGACGCTGTGTGACCTCCAAATCTTTTGGAGGAAATATGCTCCCGGCTAT 728
DB 19 CCAGGACGCTGTGTGACCTCCAAATCTTTTGGAGGAAATATGCTCCCGGCTAT 78
QY 729 CTTTATTTCAGGACGATCGTTGTATAAAGTGGCGGACTCTTCCGACAGACAGAGA 788
DB 79 CTTTATTTCAGGACGATCGTTGTATAAAGTGGCGGACTCTTCCGACAGACAGAGA 138
QY 789 GGGATTTGTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAATTTGCTTT 848
DB 139 GGGATTTGTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAATTTGCTTT 198
QY 849 GGGTGCAGCTGAGATGAGACCTTTTATTTGAATTAATTTGGAATTTGGAATTTGCAACAGCCTATGT 908
DB 199 GGGTGCAGCTGAAATGAGACCTTTTATTTGAATTAATTTGGAATTTGGAATTTGCAACAGCCTATGT 258
QY 909 CTTCTGTTATCCCAATGACGCTTCTCCCAACACCTTTGCTGACATGATGAAAGCCCTTC 968
DB 259 CTTCTGTTATCCCAATGACGCTTCTCCCAACACCTTTGCTGACATGATGAAAGCCCTTC 318
QY 969 TATGATGCCAAGCACCTTAAAGGATTTTGTCTATGATGCTTGTGTCATATAGTTGGAGG 1028
DB 319 TATGATGCCAAGCACCTTAAAGGATTTTGTCTATGATGCTTGTGTCATATAGTTGGAGG 378
QY 1029 ATGCTGTGGTCAACACAGATCATATCAGGGAATTTCTGAAGCTGTGAAATTTGTAA 1088
DB 379 ATGCTGTGGTCAACACAGATCATATCAGGGAATTTCTGAAGCTGTGAAATTTGTAA 438
QY 1089 GCCTAGAGTTCACCTGCCCTCTTTTGAAGACATATGTTACTGCTGCTGTAGAGCC 1148
DB 439 GCCTAGAGTTCACCTGCCCTCTTTTGAAGACATATGTTACTGCTGCTGTAGAGCC 498
QY 1149 CTTTCAAGGATTTGACCGTACCAACTTTTAAACATTTGGAGAGCCCTGTATTTGGAGG 1208
DB 499 CTTTCAAGGATTTGACCGTACCAACTTTTAAACATTTGGAGAGCCCTGTATTTGGAGG 558
QY 1209 ATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACTATGAAAGCCCTGTGTGTGTC 1268
DB 559 ATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACTATGAAAGCCCTGTGTGTGTC 618
QY 1269 CAAAGTGCAGGTTGGAATGGAGCCCGAGTGTGATGTCACATGATGGATGCT 1328
DB 619 CAAAGTGCAGGTTGGAATGGAGCCCGAGTGTGATGTCACATGATGGATGCT 678
QY 1329 AGATGCTCAAGTGAATGACAGATTTTGCAACTTAAATTTGCTCCGAGCCAGACATGCG 1388
DB 679 AGATGCTCAAGTGAATGACAGATTTTGCAACTTAAATTTGCTCCGAGCCAGACATGCG 738
```

```
QY 1389 AAAGGTACCTTTGTGCATCGACTCTCCAAATTTTGTGTGATTAAGCTGGTTAAAGTG 1448
DB 739 AAAGGTACCTTTGTGCATCGACTCTCC-ATTTTGTGTGATTAAGCTGGTTAAAGTG 797
QY 1449 CTGCCAAGGAAGTGCATTTGTCAATGACATTAGTCTTGAAGAGGAGAGACGACTTCTT 1508
DB 798 CTGCC-AGGAAGTGCATTTGTC-ATAGCTTAGTCTTGAAGAGGAGAGACGACTTCTT 855
QY 1509 GGAGAAGGCCAGGA 1522
DB 856 GGAGAAGGCCANGGA 869
```

#### RESULT 9

LOCUS

BU189680

DEFINITION

AGENCOURT\_7968068 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:6170551

5', mRNA sequence.

ACCESSION

BU189680

VERSION

BU189680.1 GI:22703664

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 943)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-x@mail.nih.gov](mailto:cgabs-x@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM13537 row: h column: 08

High quality sequence stop: 587.

Location/Qualifiers

1..943

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6170551"

/tissue\_type="retinoblastoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_67"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Library constructed by Life

Technologies."

#### ORIGIN

```
Query Match      20.2%; Score 791.2; DB 5; Length 943;
Best Local Similarity 96.5%; Pred. No. 1.4e-209;
Matches 830; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

QY 2155 CAAAAAATATCCCGACCTCTCAATATATTAATCAAGGACCCCTGATGAATGAATGAAA 2214
DB 1 CAAAAAATATCCCGACCTCTCAATATATTAATCAAGGACCCCTGATGAATGAATGAAA 60
QY 2215 ATTGTGTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTAAAGTCAGCC 2274
DB 61 ATTGTGTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTAAAGTCAGCC 120
QY 2275 CGGGTTATGAAGAAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAAGAAGAGAGAA 2334
DB 121 CGGGTTATGAAGAAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAAGAAGAGAGAA 180
```

QY 2335 ACCAGAGTGTCTTACGGCAGACAGTAGAAGAGAGAGACCCCTTACCAGGGCACCATCGTCTG 2394  
 Db 181 ACCAGAGTGTCTTACGGCAGACAGTAGAAGAGAGAGACCCCTTACCAGGGCACCATCGTCTG 240  
 QY 2395 GCCACTGTAAAGCGGACGTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGC 2454  
 Db 241 GCCACTGTAAAGCGGACGTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGC 300  
 QY 2455 TGCATTAATTTCCGAGTTATTTAGTTAGGAGTATGATCCATGTGTATAGATCTGAAA 2514  
 Db 301 TGCATTAATTTCCGAGTTATTTAGTTAGGAGTATGATCCATGTGTATAGATCTGAAA 360  
 QY 2515 GCTGCTCTTCCACCAAGAGAGATATTAATTTGGCTGTGCAGACTCATCTCCCTCCCTG 2574  
 Db 361 GCTGCTCTTCCACCAAGAGAGATATTAATTTGGCTGTGCAGACTCATCTCCCTCCCTG 420  
 QY 2575 GATGAATATGATTTTGTGTCACCAAGAAATGAGAGATAGTATTAAGGATTTCCATTTTG 2634  
 Db 421 GATGAATATGATTTTGTGTCACCAAGAAATGAGAGATAGTATTAAGGATTTCCATTTTG 480  
 QY 2635 ATTGGAGGAGCAACCACTTCAAAAACCCACACAGAGTTAAATAGTCCGAGATACAGT 2694  
 Db 481 ATTGGAGGAGCAACCACTTCAAAAACCCACACAGAGTTAAATAGTCCGAGATACAGT 540  
 QY 2695 GCACCTGTATCCATGTCCTGGACGCTCCAAAGAGTGTGTGTCTTCCAGCTGTTA 2754  
 Db 541 GCACCTGTATCCATGTCCTGGACGCTCCAAAGAGTGTGTGTCTTCCAGCTGTTA 600  
 QY 2755 GATGAATATGATTTTGTGTCACCAAGAAATGAGAGATAGTATTAAGGATTTCCATTTTG 2814  
 Db 601 GATGAATATGATTTTGTGTCACCAAGAAATGAGAGATAGTATTAAGGATTTCCATTTTG 660  
 QY 2815 CAGSACCATTAAGTCTCTCAAGAGAGAGAGATATTAATCCCTTAAGTCAAGCCAGAAAA 2874  
 Db 661 CAGSACCATTAAGTCTCTCAAGAGAGAGAGATATTAATCCCTTAAGTCAAGCCAGAAAA 720  
 QY 2875 AGTGGTTTCCAAATGAGTGTCTGTAACCTCACCAGTGAAGCCAGTTT-ATTGG 2933  
 Db 721 AGTGGTTTCCAAATGAGTGTCTGTAACCTCACCAGTGAAGCCAGTTT-ATTGG 780  
 QY 2934 GACCCAGTCTTTGAAGACTATGACCTGC-AGAAGCTGGTGGACTACATGACTGGAGC 2992  
 Db 781 GACCCAGTCTTTGAAGACTATGACCTGC-AGAAGCTGGTGGACTACATGACTGGAGC 840  
 QY 2993 CTTTCTTTGATGTCTGGCAG 3012  
 Db 841 CTTTCTTTGATGTCTGGCAG 860

RESULT 10  
 BX332266  
 LOCUS BX332266 961 bp mRNA linear EST 08-APR-2004  
 DEFINITION BX332266 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 cDNA clone CS0DC010Y017 5-PRIME, mRNA sequence.  
 ACCESSION BX332266  
 VERSION BX332266.2 GI:46280008  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 961)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 2, 2003 this sequence version replaced gi:30343109.  
 CONTACT Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 740.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DC010AH09QP1&c=740.f.

Location/Qualifiers

FEATURES  
 source

1. 961  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC010Y017"  
 /tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 18.3%; Score 718.4; DB 5; Length 961;  
 Best Local Similarity 92.4%; Pred. No. 3.6e-189;  
 Matches 749; Conservative 26; Mismatches 33; Indels 3; Gaps 3;  
 QY 2 GTCACTGTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGAGAGAGACTCGACA 61  
 Db 139 GTCACTGTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGAGAGAGACTCGACA 198  
 QY 62 ACATGTCACCCCGGCTCCAAAGACCTCTGCGCAACCCGAAGCTCTGGAAGAGAGAGAGAG 121  
 Db 199 ACATGTCACCCCGGCTCCAAAGACCTCTGCGCAACCCGAAGCTCTGGAAGAGAGAGAG 257  
 QY 122 ATGAGATCAATGTCATCTCTGCGAAGAGAGATTAATGCTGTGGATGGAAGAGAGAGAG 180  
 Db 258 ATGAGATCAATGTCATCTCTGCGAAGAGAGATTAATGCTGTGGATGGAAGAGAGAGAG 317  
 QY 181 ATGATCAGCGGAGAGAGCTTAACGAGAGACACTTCCGAGGTGAGGATTAAGATCAT 240  
 Db 318 ATSMTCASCGGAGAGAGCTTAACGAGAGACACTTCCGAGGTGAGGATTAAGATCAT 377  
 QY 241 GCCAGCGCTGAAAGCAACATGACATTTTAAGTATTAATGATTAATGATTAATGATTAAT 300  
 Db 378 SCASGCGCTGAAAGCAACATGACATTTTAAGTATTAATGATTAATGATTAATGATTAAT 437  
 QY 301 CAAATCCATAAGGAATACTTGTGCTGGCTGGGCGAGATATATGAAACAAATACATTTTTC 360  
 Db 438 CAAATCCATAAGGAATACTTGTGCTGGCTGGGCGAGATTAATGAAACAAATACATTTTTC 497  
 QY 361 AGCATAGTATTTGCCAAGCTGACTATGCTGCTTGAACACTTGGGCTTACCGGATGAACATG 420  
 Db 498 ACGCTCGTATTTGCCAAGCTGCTTCCGCTTTGAACACTTGGGCTTACCGGATGAACATG 557  
 QY 421 TGCTCTGCGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGAGAGAGAGAGTAAG 480  
 Db 558 TGCTCTGCGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGAGAGAGAGAGTAAG 617  
 QY 481 AGGTTTGTGGCAGGCTCTGCGTCCGACTTAATAAGACACTCTCTGTGTCCTCCATCTGTG 540  
 Db 618 AGGTTTGTGGCAGGCTCTGCGTCCGACTTAATAAGACACTCTCTGTGTCCTCCATCTGTG 677  
 QY 541 GAAAGCCCGGATTAAGGAACATCACTATTTGATGAGCTTGTGAAGAGATACCAAGAGCAG 600  
 Db 678 GAAAGCCCGGATTAAGGAACATCACTATTTGATGAGCTTGTGAAGAGATACCAAGAGCAG 737  
 QY 601 GCCAAGGAGTCTGATGCGGGGTTGATATCTTACTCAATGAACTATTTTGTACT 660  
 Db 738 GCCAAGGAGTCTGATGCGGGGTTGATATCTTACTCAATGAACTATTTTGTACT 797  
 QY 661 GCCAAGTCCAAAGCAGCTTGTGTCATCTCCAAATC-TTTTTGAGGAGAGAAATATGCTCC 719  
 Db 798 GCCAAGTCCAAAGCAGCTTGTGTCATCTCCAAATC-TTTTTGAGGAGAGAAATATGCTCC 857  
 QY 720 CCGGCTATCTTTTATTTTCAGGGAGCAGATCGTTGATTAAGAGTGGCGGAGACTCTTTCCG 779

```

Db      858  CGGCGCTATCTTTATTTTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTTCGGACA 917
      |||
Qy      780  GACAGGAGAGGATTTGTCATCAGCGTGCT 810
      |||
Db      918  GACASGAGAGGRTTTKTCATCRGCTGTCTMT 948
      |||

RESULT 11
AL598311
LOCUS      721 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFP313F151_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION  AL598311
VERSION     AL598311.1 GI:15161002
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 721)
            Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).
            EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and
            Wiemann,S.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No sl sequence available.
            This clone (DKFP313F151) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
            Location/Qualifiers
                1..721
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFP313F151"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: Sf1A; Site_2: Sf1B;
                cDNA-collection"

FEATURES
            source
            1846  CATGGGGTTTCCTTTACCATGCAATCAAGTCTGGCATGGACATGAGATAGTGAATGCT 1905
            Db      1  CATGGGGTTTCCTTTACCATGCAATCAAGTCTGGCATGGACATGAGATAGTGAATGCT 60
            Qy      1906  GGAACCTCCCTGTGTATGATATATCAATGAAGAACTCTGCAGCTCTGTGAAGATCTC 1965
            Db      61  GGAACCTCCCTGTGTATGATATATCAATGAAGAACTCTGCAGCTCTGTGAAGATCTC 120
            Qy      1966  ATCTGGAAATAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGCCCAGACCAAGGC 2025
            Db      121  ATCTGGAAATAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGCCCAGACCAAGGC 180
            Qy      2026  ACAGAGGAGGAAGAGTCAATTCAGACTGATGATGGAGAATAGCCCTGTGAGAGACGC 2085
            Db      181  ACAGAGGAGGAAGAGTCAATTCAGACTGATGATGGAGAATAGCCCTGTGAGAGACGC 240
            Qy      2086  CTTGAGTATGCCCTTGTGAAGGGCATTCGAAAAACATATATTAGGATACCTGAGGAAGCC 2145
            |||
            858  CGGCGCTATCTTTATTTTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTTCGGACA 917
            |||
            780  GACAGGAGAGGATTTGTCATCAGCGTGCT 810
            |||
            918  GACASGAGAGGRTTTKTCATCRGCTGTCTMT 948
            |||
            241  CTTGAGTATGCCCTTGTGAAGGCACTTGAAGAAACATATTATTGAGGATCTGAGGAAGCC 300
            |||
            2146  AGGTTAAACCAAAAAATATCCCGACCTCTCAATATAATTTGAAGACCCCTGATGAAT 2205
            |||
            301  AGGTTAAACCAAAAAATATCCCGACCTCTCAATATAATTTGAAGACCCCTGATGAAT 360
            |||
            2206  GGAATGAAAAATTTGTTGGTATCTTTTGGAGCTTGGAAAAATGTTTCTACCTCAGGTTATA 2265
            |||
            361  GGAATGAAAAATTTGTTGGTATCTTTTGGAGCTTGGAAAAATGTTTCTACCTCAGGTTATA 420
            |||
            2266  AAGTCAGCCCGGGTTATGAAGAAGGCTTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAA 2325
            |||
            421  AAGTCAGCCCGGGTTATGAAGAAGGCTTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAA 480
            |||
            2326  AGAGAAGAAACAGAGTGTCTTAAGCGCACAGTAGTAGAAGAGGACCTTACCAGGGCACC 2385
            |||
            481  AGAGAAGAAACAGAGTGTCTTAAGCGCACAGTAGTAGAAGAGGACCTTACCAGGGCACC 540
            |||
            2386  ATCGTGTGGCCACTGTTTAAAGCGCAGTCACACATAGGCAAGAACATAGTTGGAGTA 2445
            |||
            541  ATCGTGTGGCCACTGTTTAAAGCGCAGTCACACATAGGCAAGAACATAGTTGGAGTA 600
            |||
            2446  GTCCCTGGCTGCATAAATTTCCGAGTTATTGATTAGGAGTCATGACTCCATGTGTAAG 2505
            |||
            601  GTCCCTGGCTGCATAAATTTCCGAGTTATTGATTAGGAGTCATGACTCCATGTGTAAG 660
            |||
            2506  ATACTGAAGCTGCTCTTGACCAACAAGCAGATATTAATTTGGCCTGT-CAGGACTCATCAC 2564
            |||
            661  ATACTGAAGCTGCTCTTGACCAACAAGCAGATATTAATTTGGCCTGT-CAGGACTCATCAC 720
            |||
            2565  T 2565
            721  T 721

RESULT 12
CN414972
LOCUS      598 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION 17000418008177 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN414972
VERSION     CN414972.1 GI:47402566
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 598)
            Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
            Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
            Nat. Biotechnol. 22 (6), 707-716 (2004)
JOURNAL   15146197
PUBMED    Contact: Brandenberger R
            Regenerative Medicine
            Genon Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@genon.com
            Insert Length: 598 Std Error: 0.00.
            Location/Qualifiers
                1..598
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="embryonic stem cells, cell lines H1, H7, and
                H9"
                /clone_lib="GRN_ES"
                /note="oligo dt primed, full-length enriched cDNA library"

```

from undifferentiated hES cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## ORIGIN

Query Match 17.2%; Score 672.4; DB 7; Length 698;  
Best Local Similarity 99.9%; Pred. No. 2.6e-176;  
Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1006 GGCTTGTCATATAGTTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTT 1065  
DB 25 GGCTTGTCATATAGTTGAGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTT 84  
QY 1066 GCTGAAGCTGTGAATAATTTGAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACAT 1125  
DB 85 GCTGAAGCTGTGAATAATTTGAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACAT 144  
QY 1126 ATGTTACTGCTGCTAGAGCCCTTCAGGATTCGACCGGTACACCACTTTGTTAAACATT 1185  
DB 145 ATGTTACTGCTGCTAGAGCCCTTCAGGATTCGACCGGTACACCACTTTGTTAAACATT 204  
QY 1186 GAGAGCGCTGTAAATGTTGAGGATCAAGGAAGTTTCTTAACTCATCATGCGAGGAAC 1245  
DB 205 GAGAGCGCTGTAAATGTTGAGGATCAAGGAAGTTTCTTAACTCATCATGCGAGGAAC 264  
QY 1246 TATGAAGAAGCCTTGTCTGTTGCCAAGTGCAGTGGGAATGGAGCCAGGTGTTGGAT 1305  
DB 265 TATGAAGAAGCCTTGTCTGTTGCCAAGTGCAGTGGGAATGGAGCCAGGTGTTGGAT 324  
QY 1306 GTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1365  
DB 325 GTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 384  
QY 1366 ATTGCTTCCAGCGAGACATCGCAAGAGTACCTTTGTCATCGACTCCTCAAATTTGCT 1425  
DB 385 ATTGCTTCCAGCGAGACATCGCAAGAGTACCTTTGTCATCGACTCCTCAAATTTGCT 444  
QY 1426 GTGATTAAGCTGGTTAAAGTGTGCAAGGAGTGCATTTGTCATGATGATGATGATGATGAT 1485  
DB 445 GTGATTAAGCTGGTTAAAGTGTGCAAGGAGTGCATTTGTCATGATGATGATGATGATGAT 504  
QY 1486 AAGGAAGGAGGAGGACACTTCTTGGAAGCCAGGAGGATTTAAAGTAGTGGAGCTGCT 1545  
DB 505 AAGGAAGGAGGAGGACACTTCTTGGAAGCCAGGAGGATTTAAAGTAGTGGAGCTGCT 564  
QY 1546 ATGCTGTCATGCTTTGATGAAGAAGGACAGCAACAGAAACAGACACAAAAATCAGA 1605  
DB 565 ATGCTGTCATGCTTTGATGAAGAAGGACAGCAACAGAAACAGACACAAAAATCAGA 624  
QY 1606 GTGTGACCCGGGCTACCATCTGCTGTGAAAAAAGTGGGCTTTAATCCAAATGACATT 1665  
DB 625 GTGTGACATCCGGGCTACCATCTGCTGTGAAAAAAGTGGGCTTTAATCCAAATGACATT 684  
QY 1666 ATTTTTCACCCCTAA 1679  
DB 685 ATTTTTCACCCCTAA 698

## RESULT 13

DN280822 1161963 MARC 7BOV Bos taurus cdna 5', mRNA sequence. EST 02-MAR-2005

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Bos taurus

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

## Pecora; Bovidae; Bovinae; Bos.

## 1 (bases 1 to 715)

## Smith, T.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,

## Wray, J.E. and Keefe, J.W.

## A second set of bovine ESTs from pooled-tissue normalized libraries

## TITLE

JOURNAL  
COMMENT

Unpublished (2003)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified with  
cross\_match v0.990329.  
Plate: RLK8011 row: D column: 8  
Seq primer: GTAATACGACTCACTATAGG.  
Location/Qualifiers  
1..715  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 7BOV"  
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;  
Library made with RNA pooled from multiple tissues  
including ovary, hindbrain, uterus, and day-30 whole  
embryos."

## ORIGIN

Query Match 15.6%; Score 611.4; DB 8; Length 715;  
Best Local Similarity 92.0%; Pred. No. 3.3e-159;  
Matches 656; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 1977 AGNCCCTGAGGCCACTGAGAAGCTCTTAGTTATGCCAGACTCAAGGCAAGGAGGAA 2036  
DB 2 AGNCCCTGAGGCCACTGAGAAGCTCTTAGTTATGCCAGACTCAAGGCAAGGAGGAA 61  
QY 2037 GAAAGTCATTTCAGACTGATGAGTGGAGAAATGCGCTGTGGAAGAGCGCTTCGATATGC 2096  
DB 62 GAAAGTCATCCAGACTGATGAGTGGAGAAATGCGCTGTGGAAGAGCGCTTCGATATGC 121  
QY 2097 CCTTGTGAAGGGCATTGAAAAACATATTTAGAGTACTGTGAGAGAGCCAGGTTAAACCA 2156  
DB 122 TCTTGTGAAGGGCATTGAAAAACATATTTAGAGTACTGTGAGAGAGCCAGGTTAAACCA 181  
QY 2157 AAAAAATATCCCGACCTCTCAATATATTTAGAGAGCCCTGATGATGATGATGATGAT 2216  
DB 182 GGAATAATATCCCGACCTCTCAATATATTTAGAGAGCCCTGATGATGATGATGATGAT 241  
QY 2217 TGTGTGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTTATAAGTCAGCCCG 2276  
DB 242 TGTGTGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTTATAAGTCAGCTCG 301  
QY 2277 GGTATGAAGAAGGCTGTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAGAGAGAAC 2336  
DB 302 GGTATGAAGAAGGCTGTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAGAGAGAAC 361  
QY 2337 CAGAGTGTTAAGCGCACAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2396  
DB 362 CAAAGTGTCTTACTGGCAAAATAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421  
QY 2397 CACTGTTAAGCGCAGCTGCAGACATAGCAAGACATAGTTGGAGTAGTCTCTGCTGCTG 2456  
DB 422 TACTGTTAAGGTTGAGCTGCAGACATAGCAAGACATAGTTGGAGTAGTCTCTGCTGCTG 481  
QY 2457 CAATAATTTCCGAGTTATTGATTTAGGAGTCTAGCTCCATGTGATAAGATCTGAAAGC 2516  
DB 482 CAATAATTTCCGAGTTATTGATTTAGGAGTCTAGCTCCATGTGATAAGATCTGAAAGC 541  
QY 2517 TGCTTTGACCAAGAGAGATATATTTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2576  
DB 542 TGCTTTGACCAAGAGAGATATATTTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 601  
QY 2577 TGAATGATTTTGTGCGCTAAGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2636  
DB 602 TGAATGATTTTGTGCGCTAAGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661

Qy	2637	TGGAGGAGCAACCACTTCAAAACCCACACA-GCAGTTAAATAGTCCGAGA	2688
Db	662	CGGAGGAGTACCACCTTCCAGAACCCACAGGAGTTAAATAGCCCAAAA	714
RESULT 14			
CX228713			
LOCUS	MEM08606 Mus Musculus hematopoietic BM-HPCS cDNA library Mus		
DEFINITION	musculus cDNA 5', mRNA sequence.		
ACCESSION	CX228713		
VERSION	CX228713.1 GI:56884005		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciurognathi; Muridea; Muridae; Murinae; Mus.		
TITLE	Williams, C., Wirta, V., Richter, K., Karlsson, C., Lundberg, J. and		
	Carlsson, L.		
JOURNAL	Expressed sequence tags of cDNA clones from a hematopoietic stem		
COMMENT	cell line expressing Lhx2		
	Unpublished (2005)		
	Contact: Williams, C.		
	Molecular Biotechnology		
	Institution of Biotechnology		
	Albano University Center, KTH-Royal Institute of Technology, 106		
	91 Stockholm, Sweden		
	Tel: +46855378332.		
	Fax: +46855378481		
	Email: cecilia.williams@biotech.kth.se		
Seq primer:	M13REV.		
Location/Qualifiers	1. 758		
FEATURES	/organism="Mus musculus"		
source	/mol_type="mRNA"		
	/strain="C57BL/6-cast"		
	/db_xref="taxon:10090"		
	/sex="Male and Female"		
	/tissue type="Bone Marrow"		
	/cell type="Hematopoietic progenitor/stem cells		
	immortalized by Lhx2"		
	/cell_lines="hematopoietic stem cell-like cell line		
	BM-HPCS"		
	/dev stages="adult"		
	/clone lib="Mus Musculus hematopoietic BM-HPCS cDNA		
	library"		
	/note="Organ: Bone Marrow; Vector: pCMVSPORT6.1;		
	Preamplified custom cDNA library by Invitrogen/ResGen"		
ORIGIN			
Query Match	15.3%; Score 598.8; DB 8; Length 758;		
Best Local Similarity	87.7%; Pred. No. 1.1e-155;		
Matches	665; Conservative 0; Mismatches 92; Indels 1; Gaps 1;		
Qy	2631	GTGATTTGGAGGAGCAACCACTTCAAAACCCACACAGCAGTTAAATAGTCCGAGATA	2690
Db	1	GTGATAGGAGGAGCAACCACTTCAAGAACCCACACAGCAGTTAAATAGCACCAGTA	60
Qy	2691	CAGTGCACTGTATCCATGTCCTGGAGCGCTCCAAAGAGTGGTGGTGTCTCCAGCT	2750
Db	61	CAGCGCGCTGTGATCCACGTCCTAGATGCATCCAAGAGTGGTGGTGTCTTCAGCT	120
Qy	2751	GTATAGTAAATCTAAAGGATGAATCTTTGAGCAATCATGGAAGATATGAAGATAT	2810
Db	121	GTATAGTAAATCTGAGAGATGACTACTTTGAAGAAATACTGGAAGAGTATGAAGATAT	180
Qy	2811	TAGACAGACCACTTATGATGTCCTCAAGGAGGAGATACCTTACCTTAAGTCAAGCCAG	2870
Db	181	TAGACAGGACCATAATGATGTCCTCAAGAGAGAGAAATACGTACCCCTAAAGTCAAGCCAG	240
Qy	2871	AAAAAGTGGTTTCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCCGCTTTAT	2930
Db	241	AAAAACGGTTTCCACATTGATTTGGCTCTCTGAACCTCATCCAGTGAAGCCCGCTTTAT	300
Qy	2931	TGGGACCCAGGTCTTTGAAGACTATGACCTTGCAAGAGCTGGTGGACTACATTGACTGGAA	2990
Db	301	TGGGACCCAGGTCTTTGAGGACTACAACTTGCAAAAGCTGGTGGACTACATTGACTGGAA	360
Qy	2991	GCCTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGAT	3050
Db	361	GCCTTTCTTTGACGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGATTTCCCAAGAT	420
Qy	3051	ATTTAACGACAAAACAGTAGTGGAGAGCGCCAGGAAGGTCTACGATGATGCCACATAT	3110
Db	421	ATTTAATGACAAAGCAGTAGTGGAGAGCGCCAGGAAGGTATACATGATGCTCAGATAT	480
Qy	3111	GCTGAACACACTGATTAGTCAAAAGAACTCCGGGCGCGGGTGTGGTTCTGGCC	3170
Db	481	GCTGAACATAGTATTAGTCAAAAGAACTGCAGGCGCAGGGGTGTGGTTCTGGCC	540
Qy	3171	AGCACAGAGTATCCAGACGACATTCACCTGTAGCGAGA-GGCTGCTGTGCCCCAGGCTG	3229
Db	541	AGCACAGAGTGTCCAGATGACATCCACCTGTATGACAGGGGGGGGTGCCCGAGCCG	600
Qy	3230	CAGAGCCCATAGCCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCA	3289
Db	601	CTGAGCCCATAGCCACTTTCTATGGACTGAGCAGCAGGCTGAGAAGGACTCTCTTAGTA	660
Qy	3290	CGAGGCAATGACTGCTCTCAGACTTCATGCTTCCTGCAATCTTGGCATCTGGCATCGTACT	3349
Db	661	CAGACCCCTACCACTGCTCTCAGACTTCATGCTTCCTGCAATCTTGGCATCTTGTGACT	720
Qy	3350	ACCTGGGCTGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3387
Db	721	ACTTGGGCTGTGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	758
RESULT 15			
CX629973			
LOCUS	623 bp mRNA linear EST 11-AUG-2004		
DEFINITION	DKFZp46910922_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone		
ACCESSION	CX629973		
VERSION	CX629973.1 GI:51126053		
KEYWORDS	EST.		
SOURCE	Pongo pygmaeus (orangutan)		
ORGANISM	Pongo pygmaeus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominoidea; Pongo.		
REFERENCE	1 (bases 1 to 623)		
AUTHORS	Bahr, A., Lauber, J., Mewes, H.W., Well, B., Amid, C., Oeinger, A.,		
	Fobo, G., Han, M. and Wiemann, S.		
TITLE	Pongo pygmaeus mRNA (Bahr, A., Lauber, J., Mewes, H.W., et al.)		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: MIPS		
	MIPS		
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert Clone from S. Wiemann,		
	Molecular Genome Analysis, German Cancer Research Center (DKFZ);		
	Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen		
	(Hilden/Germany) within the cDNA sequencing consortium of the		
	German Genome Project. This clone (DKFZp46910922) is available at		
	the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in		
	Berlin, Germany. Please contact RZPD for ordering:		
	http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp46910922		
	Further information about the clone and the sequencing project is		
	available at http://mips.gsf.de/projects/cdna/.		
FEATURES	Location/Qualifiers		
source	1. .623		
	/organism="Pongo pygmaeus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9600"		



```
/clone="DKF2p46910922"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"

ORIGIN

Query Match      15.1%; Score 591.2; DB 7; Length 623;
Best Local Similarity 97.7%; Pred. No. 1.4e-153;
Matches 599; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 2562 CACTCCTTCCCTGATGAATGATTTTGTGCAAGGAATGAGAGATAGCTATAAG 2621
Db 11 CACTCCTTCTCTGGATGAATGATTTTGTGCAAGGAATGAGAGATAGCTATAAA 70

Qy 2622 GATTCCATTGTTGATGGAGGAGCAACACTTTCAAAAACCCACACAGCAGTTAAATAGC 2681
Db 71 GATTCCATTGTTGATGGAGGAGCAACACTTTCAAAAACCCACACAGCAGTTAAATAGC 130

Qy 2682 TCCGAGATACAGTGCACCTGTAATCCATGCTCGACGCGTCCAAAGAGTGTGGTGTG 2741
Db 131 TCCGAGATACAGTGCACCTGTAATCCATGCTCGACGCGTCCAAAGAGTGTGGTGTG 190

Qy 2742 TTCCACGCTGTAGATGAATCTAAAGATGAATCTTTGAGGAATCATGGAAGATA 2801
Db 191 TTCTCAGCTGTAGATGAATCTAAAGATGAATCTTTGAGGAATCATGGAAGATA 250

Qy 2802 TGAAGATTATAGACAGACCAATATGATGCTCTCAAGGAGAGGATACTTACCCTTAAG 2861
Db 251 TGAAGATTATAGACAGACCAATATGATGCTCTCAAGGAGAGGATACTTACCCTTAAG 310

Qy 2862 TCAAGCCAGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTTACCAGTGAAGCC 2921
Db 311 TCAAGCCAGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTTACCAGTGAAGCC 370

Qy 2922 CACGTTTATGGGACCCAGGCTTTTGAAGACTATGACTGCAGAGCTGTGGACTACAT 2981
Db 371 CACGTTTATGGGACCCAGGCTTTTGAAGACTATGACTGCAGAGCTGTGGACTACAT 430

Qy 2982 TGACTGGAAGCCTTTCTTTGATGCTGCGAGCTCCGGGGCAAGTACCAGATCGAGGCTT 3041
Db 431 TGACTGGAAGCCTTTCTTTGATGCTGCGAGCTCCGGGGCAAGTACCAGATCGAGGCTT 490

Qy 3042 CCCCAAGATATTTAACGACAAAAACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGC 3101
Db 491 TCCCAAGATATTTAACGACAAAAACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGC 550

Qy 3102 CCACATATGCTGAACACACTGATAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTGG 3161
Db 551 CCNAAATATGCTGAACACACTGATAGTCAAAAGAACTCCAGCCCGGGGTGTGGTGG 610

Qy 3162 GTTCTGGCCAGCA 3174
Db 611 GTTCTGGCCAGCA 623
```

Search completed: March 6, 2006, 18:54:25  
Job time : 9732 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 16:12:52 ; Search time 464 Seconds  
(without alignments)  
15013.500 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcacttgagagcagc.....ctcaaggaaatacaacctag 3919

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/ina/backfileseq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3919	100.0	3919	3	US-08-980-326-1
2	3884.8	99.1	7122	3	US-09-318-448-2
3	3884.8	99.1	7122	3	US-09-347-878-4
4	3884.8	99.1	7122	3	US-09-577-266-2
5	3883.2	99.1	7224	3	US-09-347-878-6
6	3876.2	98.9	7224	3	US-09-962-665-1
7	3876.2	98.9	7224	3	US-09-963-333-1
8	3876.2	98.9	7224	3	US-09-962-677-1
9	3851	98.3	3856	3	US-08-980-326-75
10	903	23.0	3762	3	US-09-252-991A-13448
11	903	23.0	3795	3	US-09-252-991A-13693
12	901	23.0	3750	3	US-09-343-681A-3077
13	895.8	22.9	3702	3	US-09-328-352-1578
14	701	17.9	3138	3	US-09-489-039A-1418
15	615.8	15.7	2349	3	US-09-252-991A-13893
16	274.4	7.0	1356	3	US-09-252-991A-13894
17	245.6	6.3	1830121	3	US-09-557-884-1
18	245.6	6.3	1830121	3	US-09-643-990A-1
19	245.6	6.3	1830121	3	US-10-158-865-1
20	206.8	5.3	5761	3	US-09-902-540-782
21	198.6	5.1	2304	3	US-09-902-540-7830
22	179	4.6	495	3	US-09-489-039A-1435
23	112.8	2.9	4301	3	US-09-919-891-1
24	89.6	2.3	42157	3	US-08-311-731A-126

C 25	88.2	2.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 26	87.8	2.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 27	51.4	1.3	7218	2	US-08-232-463-14	Sequence 14, Appli
C 28	41.2	1.1	1141	3	US-09-806-708B-22	Sequence 22, Appli
C 29	39.6	1.0	3586	3	US-08-847-296B-3	Sequence 3, Appli
C 30	39	1.0	3793	3	US-09-489-847-126	Sequence 126, App
C 31	39	1.0	6617	3	US-09-689-065B-1	Sequence 1, Appli
C 32	38.4	1.0	609	3	US-09-252-991A-8219	Sequence 8219, Ap
C 33	38.4	1.0	780	3	US-09-252-991A-7987	Sequence 7987, Ap
C 34	38.4	1.0	832	3	US-09-621-976-2813	Sequence 2813, Ap
C 35	38.4	1.0	1395	3	US-09-653-730-36	Sequence 36, Appli
C 36	38.4	1.0	1608	3	US-09-252-991A-8274	Sequence 8274, Ap
C 37	38.4	1.0	2055	3	US-09-252-991A-8047	Sequence 8047, Ap
C 38	38.4	1.0	3798	3	US-09-489-847-95	Sequence 95, Appli
C 39	38.4	1.0	22089	3	US-09-949-016-16890	Sequence 16890, A
C 40	38	1.0	678533	3	US-09-949-016-14577	Sequence 14577, A
C 41	38	1.0	678533	3	US-09-949-016-14578	Sequence 14578, A
C 42	37.6	1.0	289	3	US-09-007-005-17	Sequence 17, Appli
C 43	37.6	1.0	289	3	US-09-244-796-17	Sequence 17, Appli
C 44	37.4	1.0	83210	3	US-09-949-016-14209	Sequence 14209, A
C 45	37	0.9	945	3	US-09-902-540-8070	Sequence 8070, Ap

#### ALIGNMENTS

##### RESULT 1

US-08-980-326-1  
; Sequence 1, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Govette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 3919  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Other  
; LOCATION: (1)...(3919)  
; OTHER INFORMATION: Entire cloned cDNA encoding wild type methionine  
; OTHER INFORMATION: synthase.  
US-08-980-326-1

Query Match	100.0%;	Score 3919;	DB 3;	Length 3919;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3919;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGTCACTCTGAGAGACAGTCTTCTCTCCGCGCCCTCTGCGCAAGGAGAGACTCGAC	60	
Db	1	GGTCACTCTGAGAGACAGTCTTCTCTCCGCGCCCTCTGCGCAAGGAGAGACTCGAC	60	
QY	61	AACATGTCAACCGCGCTCCAGACCTGTGCGAACCCGAGAGGTCTGAAGAAAACCTTCGCG	120	
Db	61	AACATGTCAACCGCGCTCCAGACCTGTGCGAACCCGAGAGGTCTGAAGAAAACCTTCGCG	120	
QY	121	GATGAGATCAATGCCATTCTGAGAGAGATTATGCTGCTGATCGAGGATGGGACC	180	

Db 121 GATGAGTCAATGCCATTCTGCAGAGAGGATTATGGTCTGCTGATGGAGGATGGGACC 180  
Qy 181 ATGATCCAGCGGAGAGAGCTAAACGAAGAACACATTCGAGGTCAGGATTTAAAGATCAT 240  
Db 181 ATGATCCAGCGGAGAGAGCTAAACGAAGAACACATTCGAGGTCAGGATTTAAAGATCAT 240  
Qy 241 GCCAGGCGCTGAAAGGCAACATTCACATTTAAAGTATAAATCAGCCCTGATGCTATTAC 300  
Db 241 GCCAGGCGCTGAAAGGCAACATTCACATTTAAAGTATAAATCAGCCCTGATGCTATTAC 300  
Qy 301 CAAATCCATAGGAATACTTCTGGCTGGGCGACATATCATTTGAAAAAATAATCTTTAGC 360  
Db 301 CAAATCCATAGGAATACTTCTGGCTGGGCGACATATCATTTGAAAAAATAATCTTTAGC 360  
Qy 361 AGCACTAGTATGCGCCAGCTGACTATGGCCCTTGAACATCTGGCCCTACCGGATGAACATG 420  
Db 361 AGCACTAGTATGCGCCAGCTGACTATGGCCCTTGAACATCTGGCCCTACCGGATGAACATG 420  
Qy 421 TGCTCTGAGGAGTGGCCAGAAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAG 480  
Db 421 TGCTCTGAGGAGTGGCCAGAAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAG 480  
Qy 481 AGGTTTGTGCGAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCTCCATCTGTG 540  
Db 481 AGGTTTGTGCGAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCTCCATCTGTG 540  
Qy 541 GAAAGGCGGATTAAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAG 600  
Db 541 GAAAGGCGGATTAAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAG 600  
Qy 601 GCCAAAGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGAAGACTATTTTGTGATACT 660  
Db 601 GCCAAAGACTTCTGGATGGCGGGTGTGATATCTTACTCATTTGAAGACTATTTTGTGATACT 660  
Qy 661 GCCAATGCCAAGGCGGCTCTGTTGCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
Db 661 GCCAATGCCAAGGCGGCTCTGTTGCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
Qy 721 CGGCTATCTTTATTCAGGAGCGATCGTTGATTAAGTGGCGGACTCTTTCGGACAG 780  
Db 721 CGGCTATCTTTATTCAGGAGCGATCGTTGATTAAGTGGCGGACTCTTTCGGACAG 780  
Qy 781 ACAGGAGAGGATTTGTCACTCAGCGTGTCTCATGGAGAACCACTCTGATTTGGATTAAT 840  
Db 781 ACAGGAGAGGATTTGTCACTCAGCGTGTCTCATGGAGAACCACTCTGATTTGGATTAAT 840  
Qy 841 TGTGCTTTGGGTGCACTGAGATGAGACCTTTTATTTGAATTAATTTGGAATGTAACA 900  
Db 841 TGTGCTTTGGGTGCACTGAGATGAGACCTTTTATTTGAATTAATTTGGAATGTAACA 900  
Qy 901 GCCTATGTCCTGTATTCCCATGCGAGTCTTCCCAACACCTTTGGTGTGACTATGATGA 960  
Db 901 GCCTATGTCCTGTATTCCCAATGCGAGTCTTCCCAACACCTTTGGTGTGACTATGATGA 960  
Qy 961 ACGCTTCTATGATGGCCCAAGCCTAAAGATTTTGTATGGATGGCTTGTCAATATA 1020  
Db 961 ACGCTTCTATGATGGCCCAAGCCTAAAGATTTTGTATGGATGGCTTGTCAATATA 1020  
Qy 1021 GTTGGAGGATGCTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAAGCTGTGAAA 1080  
Db 1021 GTTGGAGGATGCTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAAGCTGTGAAA 1080  
Qy 1081 AATTGTAAAGCTTAGATTTCCAGCTGCTGCTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
Db 1081 AATTGTAAAGCTTAGATTTCCAGCTGCTGCTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
Qy 1141 CTAGAGCCCTTCAGGATTTGGACCGTACCAACTTTTGTAACTTTGGAGAGCGCTGTAAT 1200  
Db 1141 CTAGAGCCCTTCAGGATTTGGACCGTACCAACTTTTGTAACTTTGGAGAGCGCTGTAAT 1200  
Qy 1201 GTTGAGGATCAAGGAAGTTTGTCTAACTCATGCGAGGAACTATGAAGAGCCCTTG 1260  
Db 1201 GTTGAGGATCAAGGAAGTTTGTCTAACTCATGCGAGGAACTATGAAGAGCCCTTG 1260

Qy 1261 TGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCAGGTGTTGGATGTCAACATCGATCAT 1320  
Db 1261 TGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCAGGTGTTGGATGTCAACATCGATCAT 1320  
Qy 1321 GGCAATGCTAGATGGTCCAAAGTGCATGACAGATTTTGGCAACTTAATTTGCTTCGAGCCA 1380  
Db 1321 GGCAATGCTAGATGGTCCAAAGTGCATGACAGATTTTGGCAACTTAATTTGCTTCGAGCCA 1380  
Qy 1381 GACATCGCAAGGTACCTTTTGTGATCGACTCTCCAAATTTTGTGATTTGAAGCTGGG 1440  
Db 1381 GACATCGCAAGGTACCTTTTGTGATCGACTCTCCAAATTTTGTGATTTGAAGCTGGG 1440  
Qy 1441 TTTAAAGTGTCTCAAGGGAAGTGCATTTGCTCAATAGCAATTAGTCTGAAGGAAGAGAGGAC 1500  
Db 1441 TTTAAAGTGTCTCAAGGGAAGTGCATTTGCTCAATAGCAATTAGTCTGAAGGAAGAGAGGAC 1500  
Qy 1501 GACTTCTTGGAGAGGCGCAGGAAGATTAAAGATTTAAAGATTTGATGCTGATGCTCATGGCT 1560  
Db 1501 GACTTCTTGGAGAGGCGCAGGAAGATTAAAGATTTAAAGATTTGATGCTGATGCTCATGGCT 1560  
Qy 1561 TTTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCAACCCGGGCC 1620  
Db 1561 TTTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCAACCCGGGCC 1620  
Qy 1621 TACCATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAAT 1680  
Db 1621 TACCATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAAT 1680  
Qy 1681 ATCTTAAACCAATGGGACTGGAGGAAACAACTTTGATGCTCAATTTATTTATCCAT 1740  
Db 1681 ATCTTAAACCAATGGGACTGGAGGAAACAACTTTGATGCTCAATTTATTTATCCAT 1740  
Qy 1741 GCAACAAAAGTCAATTAAGAAAAATTTACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCAAC 1800  
Db 1741 GCAACAAAAGTCAATTAAGAAAAATTTACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCAAC 1800  
Qy 1801 TTGTCTTCTTCTTCCGAGGAAATGGAAGCAATTCGAGAAGCAATGCTGAGGTCTTCTCTT 1860  
Db 1801 TTGTCTTCTTCTTCCGAGGAAATGGAAGCAATTCGAGAAGCAATGCTGAGGTCTTCTCTT 1860  
Qy 1861 TACCATCAATCAAGTCTGGCAGCATGAGATAGTGAATGCTGGAACCTTCCCTGTG 1920  
Db 1861 TACCATCAATCAAGTCTGGCAGCATGAGATAGTGAATGCTGGAACCTTCCCTGTG 1920  
Qy 1921 TATGATGATATCCATAAGGAACTTCTGACGCTCTGGAAGATCTCATCTGGAATAAAGAC 1980  
Db 1921 TATGATGATATCCATAAGGAACTTCTGACGCTCTGGAAGATCTCATCTGGAATAAAGAC 1980  
Qy 1981 CCTGAGGCCACTGAGAGGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAA 2040  
Db 1981 CCTGAGGCCACTGAGAGGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAA 2040  
Qy 2041 GTCAATTCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTTGATGATGCCCC 2100  
Db 2041 GTCAATTCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTTGATGATGCCCC 2100  
Qy 2101 GTGAAGGCGCATTTGAAAAACATATTTATTTGAGGATACTGAGGAAGCCAGTTTAAACCAAAA 2160  
Db 2101 GTGAAGGCGCATTTGAAAAACATATTTATTTGAGGATACTGAGGAAGCCAGTTTAAACCAAAA 2160  
Qy 2161 AAATATCCCGACCTCTCAATATTAATTTGAAGGACCCCTGATGAATGAATGAATAATTTGT 2220  
Db 2161 AAATATCCCGACCTCTCAATATTAATTTGAAGGACCCCTGATGAATGAATGAATAATTTGT 2220  
Qy 2221 GGTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAAGTCAGCCGGGTT 2280  
Db 2221 GGTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAAGTCAGCCGGGTT 2280  
Qy 2281 ATGAAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAAAGAGAGAAACCAAGA 2340  
Db 2281 ATGAAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAAAGAGAGAAACCAAGA 2340



Db	405	TGATCCAGCGGGAAGCTAAACGAAGAAACATTTCCGAGGTGACGAATTTAAAGATCATG	464
Qy	242	CCAGGCGCTCAAGGCAACATGACATTTTAAAGTAACTCAGCCCTGATGCTATTACC	301
Db	465	CCAGGCGCTGAAAGGCAACATGACATTTTAAAGTAACTCAGCCCTGATGCTATTACC	524
Qy	302	AAATCCATAAGGAATACCTTGCTGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA	361
Db	525	AAATCCATAAGGAATACCTTGCTGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA	584
Qy	362	GCACTAGTATTGCCCAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT	421
Db	585	GCACTAGTATTGCCCAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT	644
Qy	422	GCTCTGCAGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	481
Db	645	GCTCTGCAGGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	704
Qy	482	GGTTTGTGGCAGGCGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG	541
Db	705	GGTTTGTGGCAGGCGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG	764
Qy	542	AAAGGCGGATATTAGGAACATCACTTTGATGAGCTTTGTTGAAGCATACCAAGACGAG	601
Db	765	AAAGGCGGATATTAGGAACATCACTTTGATGAGCTTTGTTGAAGCATACCAAGACGAG	824
Qy	602	CCAAAGGACTCTGATGCGGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG	661
Db	825	CCAAAGGACTCTGATGCGGGGTGTGATATCTTACTCATTTGAAACTATTTTGTGACTG	884
Qy	662	CCAAATGCAAGGAGCTTTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC	721
Db	885	CCAAATGCAAGGAGCTTTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC	944
Qy	722	GGCCTATCTTTATTTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGACAGA	781
Db	945	GGCCTATCTTTATTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGACAGA	1004
Qy	782	CAGGAGAGGATTTGTGCTACAGCGTGTCTCATGAGAAACCACTCTGCAATGGATTAATTT	841
Db	1005	CAGGAGAGGATTTGTGCTACAGCGTGTCTCATGAGAAACCACTCTCAATTTGGATTAATTT	1064
Qy	842	GTGCTTTGGGTGACGCTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATTTGAAATTT	901
Db	1065	GTGCTTTGGGTGACGCTGAAATGAGACCTTTTATTTGAAATTAATTTGAAATTTGAAATTT	1124
Qy	902	CCTATGCTCTGTTTATCCCAATGACGCTTTCCCAACACCTTTGGTGACTATGATGAAA	961
Db	1125	CCTATGCTCTGTTTATCCCAATGACGCTTTCCCAACACCTTTGGTGACTATGATGAAA	1184
Qy	962	CGCCTTCTATGATGGCCAAAGCACCTAAAGGATTTTGTATGATGGCTTGGTCAATATAG	1021
Db	1185	CGCCTTCTATGATGGCCAAAGCACCTAAAGGATTTTGTATGATGGCTTGGTCAATATAG	1244
Qy	1022	TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA	1081
Db	1245	TTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA	1304
Qy	1082	ATTGTAAGCTAGAGTTCCACTGCTGCACTGTTTTTGAAGGACATATGTTATCTGCTGCTC	1141
Db	1305	ATTGTAAGCTAGAGTTCCACTGCTGCACTGTTTTTGAAGGACATATGTTATCTGCTGCTC	1364
Qy	1142	TAGAGCCCTTCAGGATTTGGAACGATACCAACTTTGTTTAACTTGGAGAGCGCTGTAATG	1201
Db	1365	TAGAGCCCTTCAGGATTTGGAACGATACCAACTTTGTTTAACTTGGAGAGCGCTGTAATG	1424
Qy	1202	TTGCAAGGATCAAGGAAGTTTGTAAATCATCATGCGAGGAACTATTCAGAGAGCTTGT	1261
Db	1425	TTGCAAGGATCAAGGAAGTTTGTAAATCATCATGCGAGGAACTATTCAGAGAGCTTGT	1484
Qy	1262	GTGTTGCCAAAGTGCAGGTGGAATTTGGAGCCCGAGGTGTTGATGTCACCATGATGATG	1321

Db	1485	GTGTTGCCAAAGTGCAGGTGGAATTTGGAGCCCGAGGTGTTGGATGTCACCATGATGATG	1544
Qy	1322	GCATGCTAGATGTTCCAAAGTGCATGACACAGATTTTTCGAACTTTAAATTTGCTTCCGAGCCAG	1381
Db	1545	GCATGCTAGATGTTCCAAAGTGCATGACACAGATTTTTCGAACTTTAAATTTGCTTCCGAGCCAG	1604
Qy	1382	ACATCGCAAAAGGTACCTTTTGTGTCATCGACTCTTCCAAATTTTGTGCTGATTTGAAAGCTGGGT	1441
Db	1605	ACATCGCAAAAGGTACCTTTTGTGTCATCGACTCTTCCAAATTTTGTGCTGATTTGAAAGCTGGGT	1664
Qy	1442	TAAAGTGTCTCCAAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAAGGAGAGACG	1501
Db	1665	TAAAGTGTCTCCAAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAAGGAGAGACG	1724
Qy	1502	ACTTCTTGGAGAGAGCCAGGAAGATTTAAAGATTTGAGCTGTCTATGCTGCTCATGCTT	1561
Db	1725	ACTTCTTGGAGAGAGCCAGGAAGATTTAAAGATTTGAGCTGTCTATGCTGCTCATGCTT	1784
Qy	1562	TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCACTCCGGGCT	1621
Db	1785	TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCACTCCGGGCT	1844
Qy	1622	ACCATCTGCTTTGTAATAAACTGGGCTTTAATCCAAATGACATATTTTGTGACCCCTAATA	1681
Db	1845	ACCATCTGCTTTGTAATAAACTGGGCTTTAATCCAAATGACATATTTTGTGACCCCTAATA	1904
Qy	1682	TCCTTAACCATTTGGGACTGGAAATGAGGAAACACAACTTTGTATGCCATTAATTTTATCCATG	1741
Db	1905	TCCTTAACCATTTGGGACTGGAAATGAGGAAACACAACTTTGTATGCCATTAATTTTATCCATG	1964
Qy	1742	CAACAAAGTCAATTAAGAAACATTTACTCGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT	1801
Db	1965	CAACAAAGTCAATTAAGAAACATTTACTCGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT	2024
Qy	1802	TGTCCTTCTCCTTCGAGGAATGGAAGCAATTCGAGAAACATTCGATGCGGCTTTTCTTT	1861
Db	2025	TGTCCTTCTCCTTCGAGGAATGGAAGCAATTCGAGAAACATTCGATGCGGCTTTTCTTT	2084
Qy	1862	ACCATGCAATCAAGTCTGGCATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTGT	1921
Db	2085	ACCATGCAATCAAGTCTGGCATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTGT	2144
Qy	1922	ATGATGATATCCATAAGGAACTTCTGCAAGCTCTGTGGAAGATCTCATCTGGAATAAGAC	1981
Db	2145	ATGATGATATCCATAAGGAACTTCTGCAAGCTCTGTGGAAGATCTCATCTGGAATAAGAC	2204
Qy	1982	CTGAGGCCACTGAGAACTCTTACGTTATGCCAGACTCAAGGACACAGGAGGAGAAAG	2041
Db	2205	CTGAGGCCACTGAGAACTCTTACGTTATGCCAGACTCAAGGACACAGGAGGAGAAAG	2264
Qy	2042	TCATTTCAAGCTGATGAGTGGAGAAATGCGCTGTCGAAAGACGCTTGTGATGCTCCCTTG	2101
Db	2265	TCATTTCAAGCTGATGAGTGGAGAAATGCGCTGTCGAAAGACGCTTGTGATGCTCCCTTG	2324
Qy	2102	TGAAGGCGCATTTGAAAAACATATTTTGAAGGATCTGAGGAAGCCAGGTTTAAACCAAAAA	2161
Db	2325	TGAAGGCGCATTTGAAAAACATATTTTGAAGGATCTGAGGAAGCCAGGTTTAAACCAAAAA	2384
Qy	2162	AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2221
Db	2385	AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2444
Qy	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTTACCTCAGGTTTAAAGTCAAGCCCGGGTTA	2281
Db	2445	GTGATCTTTTGGAGCTGGAAAAATGTTTCTTACCTCAGGTTTAAAGTCAAGCCCGGGTTA	2504
Qy	2282	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTATGAAAAAAGAGAGAAACACAGAG	2341
Db	2505	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTATGAAAAAAGAGAGAAACACAGAG	2564
Qy	2342	TGCTTAAAGGCAAGTGAAGAGAGGACCTTTTACCGGGGACCATTCGTTGCTGGCCACTG	2401
Db	2565	TGCTTAAAGGCAAGTGAAGAGAGGACCTTTTACCGGGGACCATTCGTTGCTGGCCACTG	2624



```
QY 2402 TTAAGGCGACGTGCAACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTGGCTGCAATA 2461
Db 2625 TTAAGGCGACGTGCAACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTGGCTGCAATA 2684
QY 2462 ATTTCCGAGTTATTGATTAGGAGTCATGACTCCCATGTGTAAGATAGTGAAGAGCTGCTC 2521
Db 2685 ATTTCCGAGTTATTGATTAGGAGTCATGACTCCCATGTGTAAGATAGTGAAGAGCTGCTC 2744
QY 2522 TTGACCAAAAGCAGATATAATTGGCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2581
Db 2745 TTGACCAAAAGCAGATATAATTGGCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2804
QY 2582 TGATTTTGTGGCCAAAGAAATGAGAGATTAGCTATAGGATTTCCATTGTTGATTGGAG 2641
Db 2805 TGATTTTGTGGCCAAAGAAATGAGAGATTAGCTATAGGATTTCCATTGTTGATTGGAG 2864
QY 2642 GAGCAACCACTTCAAAAGCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701
Db 2865 GAGCAACCACTTCAAAAGCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2924
QY 2702 TAATCCATGTCTGAGCGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 2761
Db 2925 TAATCCATGTCTGAGCGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 2984
QY 2762 ATCTAAGGATGAATACCTTTGAGGAATCATGGAAGAAATATGAAGATATTAAGACGACC 2821
Db 2985 ATCTAAGGATGAATACCTTTGAGGAATCATGGAAGAAATATGAAGATATTAAGACGACC 3044
QY 2822 ATATGAGTCTCTCAAGGAGAGGAGTACTTACCCTTAAGTCAAGCCAGAAAAGTGGTT 2881
Db 3045 ATATGAGTCTCTCAAGGAGAGGAGTACTTACCCTTAAGTCAAGCCAGAAAAGTGGTT 3104
QY 2882 TCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACAGCTTTATTGGGACCCAGG 2941
Db 3105 TCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACAGCTTTATTGGGACCCAGG 3164
QY 2942 TCTTTGAGACTATGACTCGAGAGCTGTGGTGGACTACATTGACTGGAAGCCTTCTTTTG 3001
Db 3165 TCTTTGAGACTATGACTCGAGAGCTGTGGTGGACTACATTGACTGGAAGCCTTCTTTTG 3224
QY 3002 ATGTCTGCGAGCTCGGGGCAAGTACCGATCGAGGCTTCCCAAGATATTTAACGACA 3061
Db 3225 ATGTCTGCGAGCTCGGGGCAAGTACCGATCGAGGCTTCCCAAGATATTTAACGACA 3284
QY 3062 AAACAGTAGTGGAGAGGCGCAGGAGGCTACAGATGATGCCCAACAATATGCTGAACACAC 3121
Db 3285 AAACAGTAGTGGAGAGGCGCAGGAGGCTACAGATGATGCCCAACAATATGCTGAACACAC 3344
QY 3122 TGATTAGTCAAAAGAAACTCCGGGCGGGGGTGTGGTGGTCTTGCCACGACAGAGTA 3181
Db 3345 TGATTAGTCAAAAGAAACTCCGGGCGGGGGTGTGGTGGTCTTGCCACGACAGAGTA 3404
QY 3182 TCCAGACGACATTCACCTGTACGAGAGGCTGTGTGCCCCAGGCTGCAGAGCCCATAG 3241
Db 3405 TCCAGACGACATTCACCTGTACGAGAGGCTGTGTGCCCCAGGCTGCAGAGCCCATAG 3464
QY 3242 CCACCTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGAGCCATACT 3301
Db 3465 CCACCTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGAGCCATACT 3524
QY 3302 ACTGCTCTCAGACTTCACTGCTCCCTTGCATTCTTGGCATCCGTGACTTACCTGGGCTGT 3361
Db 3525 ACTGCTCTCAGACTTCACTGCTCCCTTGCATTCTTGGCATCCGTGACTTACCTGGGCTGT 3584
QY 3362 TTGCGGCTGCTGCTTTGGGTAGAGAGCTGAGCAGGCTCTATGAGATGATGGTGAAG 3421
Db 3585 TTGCGGCTGCTGCTTTGGGTAGAGAGCTGAGCAGGCTCTATGAGATGATGGTGAAG 3644
QY 3422 ACTACAGCAGCATCATGTGTCAAGGCGCTGGGGGACCGGCTGGCAGAGCCCTTTTCAGAG 3481
Db 3645 ACTACAGCAGCATCATGTGTCAAGGCGCTGGGGGACCGGCTGGCAGAGCCCTTTTCAGAG 3704
```

```
QY 3482 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGGCTACTCTGGCAGTGAAGAGCTGGACG 3541
Db 3705 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGGCTACTCTGGCAGTGAAGAGCTGGACG 3764
QY 3542 TCGCAGACCTTCCGAAGGTTCCGCTACAAGGCGCATCCGCCGGCTCTCTGGCTACCCAGCC 3601
Db 3765 TCGCAGACCTTCCGAGGCTCGGCTACAAGGCGCATCCGCCGGCTCTCTGGCTACCCAGCC 3824
QY 3602 AGCCCGACCAACCCGAGAGCTCACCATGTGGAGATCTCGCAGACATCGAGAGTCTACAG 3661
Db 3825 AGCCCGACCAACCCGAGAGCTCACCATGTGGAGATCTCGCAGACATCGAGAGTCTACAG 3884
QY 3662 GCATTAGTTAAACAGAGATCATTAGCATGCGACCTCTTCAGCAGTCTCAGGGCTCTACT 3721
Db 3885 GCATTAGTTAAACAGAGATCATTAGCATGCGACCTCTTCAGCAGTCTCAGGGCTCTACT 3944
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 3781
Db 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 4004
QY 3782 ATTATGCAATTGAGAAAGAACATATCTGTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841
Db 4005 ATTATGCAATTGAGAAAGAACATATCTGTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4064
QY 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTGGCCCTTTTATCTTGATGATCCT 3901
Db 4065 TGGGATATGATACAGACTAACT--TTTTTTTTTTTTTTTGGCCCTTTTATCTTGATGATCCT 4122
QY 3902 CAAGGAAATCAACCTAG 3919
Db 4123 CAAGGAAATCAACCTAG 4140
```

## RESULT 3

```
US-09-347-878-4
; Sequence 4, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 7122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (287)..(4084)
; OTHER INFORMATION: Human methionine synthase cDNA
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U75743/GenBank
US-09-347-878-4
```

```
Query Match 99.1%; Score 3884.8; DB 3; Length 7122;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3904; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 2 GTCACTGTGGAGAGCAGCTCTTCTCTGCCGCGCCCTCTGCGCAAGAGGAGAGTCTGACA 61
Db 225 GTCACTGTGGAGAGCAGCTCTTCTCTGCCGCGCCCTCTGCGCAAGAGGAGAGTCTGACA 284
QY 62 ACATGTCAACCCGCGCTCCAAAGACTGTGCGCAACCCGAGAGTCTGAGAGAAACCTCGGG 121
Db 285 ACATGTCAACCCGCGCTCCAAAGACTGTGCGCAACCCGAGAGTCTGAGAGAAACCTCGGG 344
QY 122 ATGAGTCAATGCCATCTCTCAGAGAGGATTTGTTGCTGGATGGAGGATGGGACCA 181
Db 345 ATGAGTCAATGCCATCTCTCAGAGAGGATTTGTTGCTGGATGGAGGATGGGACCA 404
```

QY 182 TGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTGAGGAAATTTAAAGATCATG 241  
DB 405 TGAATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTGAGGAAATTTAAAGATCATG 464  
QY 242 CCAGCCCGCTGAAGGCAACAAATGACATTTAAAGTATTAACCTCAGCCTGATGTCATTTACC 301  
DB 465 CCAGCCCGCTGAAGGCAACAAATGACATTTAAAGTATTAACCTCAGCCTGATGTCATTTACC 524  
QY 302 AAATCCATAAGGAATACCTTGGCTGGGCGAGATATCATTTGAACAAATACTTTTAGCA 361  
DB 525 AAATCCATAAGGAATACCTTGGCTGGGCGAGATATCATTTGAACAAATACTTTTAGCA 584  
QY 362 GCACTAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 421  
DB 585 GCACTAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 644  
QY 422 GCTCTGCAAGGATGGCCAGAAAGCTGCCGAGGAGGTAACTCTCAGACAGGAATTTAAGA 481  
DB 645 GCTCTGCAAGGATGGCCAGAAAGCTGCCGAGGAGGTAACTCTCAGACAGGAATTTAAGA 704  
QY 482 GGTGTGTGGCAGGGGCTCTGGGTCGACTAATTAAGACACTCTCTGTGTCCCATCTCTGG 541  
DB 705 GGTGTGTGGCAGGGGCTCTGGGTCGACTAATTAAGACACTCTCTGTGTCCCATCTCTGG 764  
QY 542 AAAGCCCGGATTAAGGAACATCACTATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 601  
DB 765 AAAGCCCGGATTAAGGAACATCACTATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 824  
QY 602 CCAAAGGACTTCTGGATCGCGGGTGTGATCTCTTACTCAATTTGAACATTTTGTGATCTG 661  
DB 825 CCAAAGGACTTCTGGATCGCGGGTGTGATCTCTTACTCAATTTGAACATTTTGTGATCTG 884  
QY 662 CCAATGCCAAGCGCCTGTGTGACCTCCAAATCTTTTGAAGGAGAAATGCTCCCC 721  
DB 885 CCAATGCCAAGCGCCTGTGTGACCTCCAAATCTTTTGAAGGAGAAATGCTCCCC 944  
QY 722 GGCCTATCTTTATTTTCAAGGACGATCGTTGATTAAGGTGGCGGACTCTTTCCGACAGA 781  
DB 945 GGCCTATCTTTATTTTCAAGGACGATCGTTGATTAAGGTGGCGGACTCTTTCCGACAGA 1004  
QY 782 CAGGAGAGGGAATTTGTCTATCAGCGGTCTCATGAGGAAACCACTCTGCAATTTGATTAAT 841  
DB 1005 CAGGAGAGGGAATTTGTCTATCAGCGGTCTCATGAGGAAACCACTCTTACATTTGATTAAT 1064  
QY 842 GTGCTTTGGGTGCAGCTGAGATGACCTTTTATTTGAATTAATTTGGAATTTGTAACAG 901  
DB 1065 GTGCTTTGGGTGCAGCTGAAATGAGACCTTTTATTTGAATTAATTTGGAATTTGTAACAG 1124  
QY 902 CCTATGCTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTGGTGACATGATGAAA 961  
DB 1125 CCTATGCTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTGGTGACATGATGAAA 1184  
QY 962 CGCCTTCTATGATGCCAAGCACTTAAAGGATTTTGTATGATGGCTTGGTGCATATATAG 1021  
DB 1185 CGCCTTCTATGATGCCAAGCACTTAAAGGATTTTGTCTATGATGGCTTGGTGCATATATAG 1244  
QY 1022 TTGAGGAGTCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1081  
DB 1245 TTGAGGAGTCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAGCTTAGAGTTCCACTGCGACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141  
DB 1305 ATTGTAAGCTTAGAGTTCCACTGCGACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1364  
QY 1142 TAGAGCCCTTCAAGGATTTGACCGGTACACCACTTTTGAACATTTGAGAGCGCTGTAAATG 1201  
DB 1365 TAGAGCCCTTCAAGGATTTGACCGGTACACCACTTTTGAACATTTGAGAGCGCTGTAAATG 1424  
QY 1202 TTGAGGAGTCAAGGAGTTTGTCTAACTCATGTCAGGAGAACTTGAAGAGCCTTGT 1261  
DB 1425 TTGAGGAGTCAAGGAGTTTGTCTAACTCATGTCAGGAGAACTTGAAGAGCCTTGT 1484  
QY 1262 GTGTTGCCAAAGTGCAGGTGGGAAATGGGAGCCAGGTGTGATGTCACATGAGTATGATG 1321

DB 1485 GTGTTGCCAAAGTGCAGGTGGGAAATGGGAGCCAGGTGTGATGTCAAATGATGATG 1544  
QY 1322 GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTGCACACTTAAATTTGCTTTCCGAGCCAG 1381  
DB 1545 GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTGCACACTTAAATTTGCTTTCCGAGCCAG 1604  
QY 1382 ACATGCAAAAGGTACCTTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1441  
DB 1605 ACATGCAAAAGGTACCTTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1664  
QY 1442 TAAAGTCTGCCAAGGAGTGCATTTGATCAATAGCTTAAAGTATGCTGAAAGGAGGAGGACG 1501  
DB 1665 TAAAGTCTGCCAAGGAGTGCATTTGATCAATAGCTTAAAGTATGCTCAAGGAGGAGGAGACG 1724  
QY 1502 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGAGTCTGCTATGGTGTGATGCTTT 1561  
DB 1725 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGAGTCTGCTATGGTGTGATGCTTT 1784  
QY 1562 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCAACCGGCGCT 1621  
DB 1785 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCAACCGGCGCT 1844  
QY 1622 ACCATCTGCTTGTCAAAAAAATCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA 1681  
DB 1845 ACCATCTGCTTGTCAAAAAAATCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA 1904  
QY 1682 TCCTAACCATTTGGGACTGGGAAATGGAGGAAACAACTTGTATGCCATTAATTTTATCCATG 1741  
DB 1905 TCCTAACCATTTGGGACTGGGAAATGGAGGAAACAACTTGTATGCCATTAATTTTATCCATG 1964  
QY 1742 CAACAAAGTCAATTAAGAAACATTAATCTGGAGCCAGAAATGAGTGGAGGCTTTTCCAACT 1801  
DB 1965 CAACAAAGTCAATTAAGAAACATTAATCTGGAGCCAGAAATGAGTGGAGGCTTTTCCAACT 2024  
QY 1802 TGTCTTCTCTCTTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGCAATGGGCTTTTCTTT 1861  
DB 2025 TGTCTTCTCTCTTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGCAATGGGCTTTTCTTT 2084  
QY 1862 ACCATGCAATCAAGTCTGGCATGGGACATGGAGATGAGTGAATGCTGGAAACCTCCCTGTG 1921  
DB 2085 ACCATGCAATCAAGTCTGGCATGGGACATGGGATGAGTGAATGCTGGAAACCTCCCTGTG 2144  
QY 1922 ATGATGATATCCATAAGGAACTTCTGCAAGCTCTGTGAAGATCTCATCTGGAATTAAGACC 1981  
DB 2145 ATGATGATATCCATAAGGAACTTCTGCAAGCTCTGTGAAGATCTCATCTGGAATTAAGACC 2204  
QY 1982 CTGAGGCCACTGGAAGCTCTTACGTTATGCCAGACTCAAGGCAACAGGAGGGAAGAAAG 2041  
DB 2205 CTGAGGCCACTGGAAGCTCTTACGTTATGCCAGACTCAAGGCAACAGGAGGGAAGAAAG 2264  
QY 2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGACCGCTTGTGATGATGCCCTTG 2101  
DB 2265 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGACCGCTTGTGATGATGCCCTTG 2324  
QY 2102 TGAAGGCAATTTGAAAAAATATTTAGGATATCTGAGGAGCCAGGTTTAAACCAAAAA 2161  
DB 2325 TGAAGGCAATTTGAAAAAATATTTAGGATATCTGAGGAGCCAGGTTTAAACCAAAAA 2384  
QY 2162 AATAATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
DB 2385 AATAATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2444  
QY 2222 GTGATCTTTTGGAGCTGGGAAAAATGTTTCTACCTCAGGTTATTAAGTCAGCCCGGTTA 2281  
DB 2445 GTGATCTTTTGGAGCTGGGAAAAATGTTTCTACCTCAGGTTATTAAGTCAGCCCGGTTA 2504  
QY 2282 TGAAGAGGCTTGTGGCCACCTTATCCCTTTTCAAGAAAAAAGAGAGAAACACAGAG 2341  
DB 2505 TGAAGAGGCTTGTGGCCACCTTATCCCTTTTCAAGAAAAAAGAGAGAAACACAGAG 2564  
QY 2342 TGCTTAAACGCGACAGTAGAAGAGAGGACCTTATACAGGCGCACCATGCTGTGCGCACCTG 2401



Db 405 TGAATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 464  
Qy 242 CCAGGCGGCTGAAGGCAACATAGACATTTAAAGTATTAACCTCAGGCTGATGTCAATTTACC 301  
Db 465 CCAGGCGGCTGAAGGCAACATAGACATTTAAAGTATTAACCTCAGGCTGATGTCAATTTACC 524  
Qy 302 AAATCCATAAGGAATACCTTGTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTTAGCA 361  
Db 525 AAATCCATAAGGAATACCTTGTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTTAGCA 584  
Qy 362 GCATCTAGTATTTGCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 421  
Db 585 GCATCTAGTATTTGCCAAGCTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGT 644  
Qy 422 GCTCTGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 481  
Db 645 GCTCTGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGA 704  
Qy 482 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAAATAAGACACTCTGTGTGCCCATCTGTGG 541  
Db 705 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAAATAAGACACTCTGTGTGCCCATCTGTGG 764  
Qy 542 AAAGGCCGGATTAAGGAACATCAATTTGATGAGCTTTGTGAAGCATACCAAGACGAGG 601  
Db 765 AAAGGCCGGATTAAGGAACATCAATTTGATGAGCTTTGTGAAGCATACCAAGACGAGG 824  
Qy 602 CCAAGGAGCTCTGGATGGCGGGTGTATATCTTACTCATTTGAACACTATTTTGTACTG 661  
Db 825 CCAAGGAGCTCTGGATGGCGGGTGTATATCTTACTCATTTGAACACTATTTTGTACTG 884  
Qy 662 CCAATGCCAAGCGACCTTTGTTCACCTCCAAATCTTTTGGAGAGAAATATGTCCTCC 721  
Db 885 CCAATGCCAAGCGACCTTTGTTCACCTCCAAATCTTTTGGAGAGAAATATGTCCTCC 944  
Qy 722 GGCCTATCTTTATTTTCAGGAGCATCGTTGTGATAAAGTGGCGGACTCTTTCCCGACAGA 781  
Db 945 GGCCTATCTTTATTTTCAGGAGCATCGTTGTGATAAAGTGGCGGACTCTTTCCCGACAGA 1004  
Qy 782 CAGGAGAGGATTTGTCTATCAGCTGTCTCATGGAGAACCACTCTGCAATGGATTTAAAT 841  
Db 1005 CAGGAGAGGATTTGTCTATCAGCTGTCTCATGGAGAACCACTCTACATTTGGATTTAAAT 1064  
Qy 842 GTGCTTTGGGTGACGTGAGATGACACTTTTATTTGAATAATTTGGAATAATGTAACAG 901  
Db 1065 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTTGAATAATTTGGAATAATGTAACAG 1124  
Qy 902 CCTATGCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTTGGTGACTATGATGAAA 961  
Db 1125 CCTATGCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTTGGTGACTATGATGAAA 1184  
Qy 962 GGCCTTTCTATGATGGCCAGACCTAAAGGATTTTGTATGGATGGCTTGGTCAATATAG 1021  
Db 1185 GGCCTTTCTATGATGGCCAGACCTAAAGGATTTTGTATGGATGGCTTGGTCAATATAG 1244  
Qy 1022 TTGAGGAGTGTGTGGGTCAACACAGATCATATCAGGAGAAATTCGCAAGCTGTGAAA 1081  
Db 1245 TTGAGGAGTGTGTGGGTCAACACAGATCATATCAGGAGAAATTCGCAAGCTGTGAAA 1304  
Qy 1082 ATTGTAAAGCTAGAGTTCCACCTGCCACTGTCTTTTGAAGGACATATGTTACTGTGGTC 1141  
Db 1305 ATTGTAAAGCTAGAGTTCCACCTGCCACTGTCTTTTGAAGGACATATGTTACTGTGGTC 1364  
Qy 1142 TAGAGCCCTTAGAGATGGACCGTACACCACTTTTGAAGGAGCGCTGTAATG 1201  
Db 1365 TAGAGCCCTTAGAGATGGACCGTACACCACTTTTGAAGGAGCGCTGTAATG 1424  
Qy 1202 TTGAGGAGTCAAGGAGTTTGTCTAACTCATCATCGGAGAACTATGAGAGGCTGTGT 1261  
Db 1425 TTGAGGAGTCAAGGAGTTTGTCTAACTCATCATCGGAGAACTATGAGAGGCTGTGT 1484  
Qy 1262 GTGTTGCCAAGTGCAGGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATG 1321

Db 1485 GTGTTGCCAAGTGCAGGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATG 1544  
Qy 1322 GCATGCTAGATGTCCTCAAGTGCATGACACAGATTTTTCGAATTTTTCCTCCAGGACG 1381  
Db 1545 GCATGCTAGATGTCCTCAAGTGCATGACACAGATTTTTCGAATTTTTCCTCCAGGACG 1604  
Qy 1382 ACATCGCAAAAGGTACCTTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGGT 1441  
Db 1605 ACATCGCAAAAGGTACCTTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGGT 1664  
Qy 1442 TAAAGTCTGCTCCAAAGGAGTGCATTTGTCTCAATAGCATTTAGTCTGGAAGGAGGAGAGACG 1501  
Db 1665 TAAAGTCTGCTCCAAAGGAGTGCATTTGTCTCAATAGCATTTAGTCTGGAAGGAGGAGAGACG 1724  
Qy 1502 ACTTCTTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGCTGTCTATGCTGCTCATGCTT 1561  
Db 1725 ACTTCTTTGGAGAGGCGCAGGAAGATTTAAAGTATGAGCTGTCTATGCTGCTCATGCTT 1784  
Qy 1562 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGCGCT 1621  
Db 1785 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGCGCT 1844  
Qy 1622 ACCATCTGCTTTGGAAGGAGGCTGGGCTTTTAAATCCAAATGACATTTATTTTGAACCCCTAATA 1681  
Db 1845 ACCATCTGCTTTGGAAGGAGGCTGGGCTTTTAAATCCAAATGACATTTATTTTGAACCCCTAATA 1904  
Qy 1682 TCCTTAACCATTTGGGACTGGAAATGGAGGAAACAACTTTGTATGCCATTTATTTTATCCATG 1741  
Db 1905 TCCTTAACCATTTGGGACTGGAAATGGAGGAAACAACTTTGTATGCCATTTATTTTATCCATG 1964  
Qy 1742 CAACAAAGTCATTTAAAGAAACATTTACTCTGGAGCCAGAAATAAGTGGAGGCTCTTTTCCAACT 1801  
Db 1965 CAACAAAGTCATTTAAAGAAACATTTACTCTGGAGCCAGAAATAAGTGGAGGCTCTTTTCCAACT 2024  
Qy 1802 TGTCTCTCTCTCCCGAGGAAATGGAGCCATTTGAGAGCAATGCAATGCAATGCGGCTTTTCTTT 1861  
Db 2025 TGTCTCTCTCTCCCGAGGAAATGGAGCCATTTGAGAGCAATGCAATGCAATGCGGCTTTTCTTT 2084  
Qy 1862 ACCATGCAATTCAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAGACCTCCCTGTGT 1921  
Db 2085 ACCATGCAATTCAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAGACCTCCCTGTGT 2144  
Qy 1922 ATGATGATATCCATAAGGAACTTTCTGCAAGCTCTGTGAGAGATCTCATCTGGAATAAGAGACC 1981  
Db 2145 ATGATGATATCCATAAGGAACTTTCTGCAAGCTCTGTGAGAGATCTCATCTGGAATAAGAGACC 2204  
Qy 1982 CTGAGGCGCATGAGAAAGCTCTTACGTTATGCCAGACTCAAGGACACAGGAGGAGAGAAAG 2041  
Db 2205 CTGAGGCGCATGAGAAAGCTCTTACGTTATGCCAGACTCAAGGACACAGGAGGAGAGAAAG 2264  
Qy 2042 TCATTTGAGCTGATGAGTGGAGAAATGGCCCTGTGGAAGAACCGCTTGGATGATGCCCTTG 2101  
Db 2265 TCATTTGAGCTGATGAGTGGAGAAATGGCCCTGTGGAAGAACCGCTTGGATGATGCCCTTG 2324  
Qy 2102 TGAAGGCGCATTTGAAAAACATATTTATGAGGATATCTGAGGAGCCAGGTTTAAACCAAAAA 2161  
Db 2325 TGAAGGCGCATTTGAAAAACATATTTATGAGGATATCTGAGGAGCCAGGTTTAAACCAAAAA 2384  
Qy 2162 AATATCCCGACCTCTCAATATAATGAAAGACCCCTGATGAATGGAATGGAATTTGTGT 2221  
Db 2385 AATATCCCGACCTCTCAATATAATGAAAGACCCCTGATGAATGGAATGGAATTTGTGT 2444  
Qy 2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAGCTCAGCCCGGGTTA 2281  
Db 2445 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAGCTCAGCCCGGGTTA 2504  
Qy 2282 TGAAGAGGCTGTTGGGCCACCTTATCCCTTTTATGAAAAAAGAGAGAGAAACCAAGAG 2341  
Db 2505 TGAAGAGGCTGTTGGGCCACCTTATCCCTTTTATGAAAAAAGAGAGAGAAACCAAGAG 2564  
Qy 2342 TGTCTTAAACGCGACATGTAAGAGAGAGAGACCTTTTACAGGAGCCCATGCTGTGCGCACTG 2401  
Db 2565 TGTCTTAAACGCGACATGTAAGAGAGAGAGACCTTTTACAGGAGCCCATGCTGTGCGCACTG 2624

2402 TTAAGGCGAGCTGCAAGACATAGGCAAGAACATAGTTGAGTAGTCTCTTGGCTGCAATA 2461  
Db TTAAGGCGAGCTGCAAGACATAGGCAAGAACATAGTTGAGTAGTCTCTTGGCTGCAATA 2464  
2462 ATTTCCGAGTATTGATTAGGAGTCATGCTCCATGTAAGATAGATCTGAAGCTGCTC 2521  
Db ATTTCCGAGTATTGATTAGGAGTCATGCTCCATGTAAGATAGATCTGAAGCTGCTC 2744  
2522 TTGACCAAAAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTGGATGAAA 2581  
Db TTGACCAAAAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTGGATGAAA 2804  
2582 TGATTTTGTGTCAGAGAAATGAGAGATTAGCTATAAGATTTCCATTGTTGATGGAG 2641  
Db TGATTTTGTGTCAGAGAAATGAGAGATTAGCTATAAGATTTCCATTGTTGATGGAG 2864  
2642 GAGCAACACATTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701  
Db GAGCAACACATTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2824  
2702 TAATCCATGTCTCGGACGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 2761  
Db TAATCCATGTCTCGGACGCGTCCAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAAA 2984  
2762 ATCTAAAGGATGAATATCTTTGAGGAATCATGGAAGATATGAAGATATTAGACAGACC 2821  
Db ATCTAAAGGATGAATATCTTTGAGGAATCATGGAAGATATGAAGATATTAGACAGACC 3044  
2822 ATTATGAGTCTCTCAAGGAGGAGATACCTTACCCTTAAGTCAAGCCAGAAAGTGGTT 2881  
Db ATTATGAGTCTCTCAAGGAGGAGATACCTTACCCTTAAGTCAAGCCAGAAAGTGGTT 3104  
2882 TCCTAAATGGATTTGGCTGTCTGAACCTCAACCCAGTGAAGCCACCGCTTTATTGGGACCCAGG 2941  
Db TCCTAAATGGATTTGGCTGTCTGAACCTCAACCCAGTGAAGCCACCGCTTTATTGGGACCCAGG 3164  
2942 TCTTTGAAGATATGACTGACAGAACTGTGGTGTGACTACATGACTGAAGCCTTTCTTTG 3001  
Db TCTTTGAAGATATGACTGACAGAACTGTGGTGTGACTACATGACTGAAGCCTTTCTTTG 3224  
3002 ATGTCTGGGAGCTCCGGGCGAAGTACCAGGATCCAGGCTTCCCAAGATATTTAAGACA 3061  
Db ATGTCTGGGAGCTCCGGGCGAAGTACCAGGATCCAGGCTTCCCAAGATATTTAAGACA 3284  
3062 AAACAGTAGGTGGAGAGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3121  
Db AAACAGTAGGTGGAGAGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3344  
3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGGTGGTGTCTGCGCCAGCACAGAGTA 3181  
Db TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGGTGGTGTCTGCGCCAGCACAGAGTA 3404  
3182 TCCAGAGCAGATTCACCTGTACCGAGGCTGTCTGCCCCAGGCTGCGAGGCCCATAG 3241  
Db TCCAGAGCAGATTCACCTGTACCGAGGCTGTCTGCCCCAGGCTGCGAGGCCCATAG 3464  
3242 CCACCTTTCTATGGTGAAGCAACAGCTGAGAGGACTCTGCCAGCAGCAGGACCACT 3301  
Db CCACCTTTCTATGGTGAAGCAACAGCTGAGAGGACTCTGCCAGCAGCAGGACCACT 3524  
3302 ACTGCTCTCAGACTTCATGCTCCCTTTCATTTCTGGCATCCGCTGACTCTGGGCTGT 3361  
Db ACTGCTCTCAGACTTCATGCTCCCTTTCATTTCTGGCATCCGCTGACTCTGGGCTGT 3584  
3362 TTGCGGTGCTGCTGTTGGGGTAGAAGCTGAGCAAGGCTGATGAGGATGATGGTGAAG 3421  
Db TTGCGGTGCTGCTGTTGGGGTAGAAGCTGAGCAAGGCTGATGAGGATGATGGTGAAG 3644  
3422 ACTACAGCAGATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAG 3481  
Db ACTACAGCAGATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAG 3704

3482 AGCTTCATGAAGAGTTCCCGGAGAACTGTGGCCCTACTGTGCGAGTGAAGCTGGAAG 3541  
Db AGCTTCATGAAGAGTTCCCGGAGAACTGTGGCCCTACTGTGCGAGTGAAGCTGGAAG 3764  
3542 TCGCAGACCTGCGAAGGTTGCGGTACAAGGGCATCCGCGCGCTCTCTGGCTACCCAGCC 3601  
Db TCGCAGACCTGCGAAGGTTGCGGTACAAGGGCATCCGCGCGCTCTCTGGCTACCCAGCC 3824  
3602 AGCCCAACCAACCGGAGAGCTCACCATGTGGAGATCTGCGAGACATCGAGAGCTCTACAG 3661  
Db AGCCCAACCAACCGGAGAGCTCACCATGTGGAGATCTGCGAGACATCGAGAGCTCTACAG 3884  
3662 GCATTAGTGAAGAGATCAATAGCAATGGCACTGCTTCCAGCTCTCAGGCTCTACT 3721  
Db GCATTAGTGAAGAGATCAATAGCAATGGCACTGCTTCCAGCTCTCAGGCTCTACT 3944  
3722 TCTCCAAATTTGAAGTCCAAATATTTTGTCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 3781  
Db TCTCCAAATTTGAAGTCCAAATATTTTGTCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 4004  
3782 ATTATGATTTGAGGAAGAACATATCTGTGCTGAGGTGAGAAATGGCTTGGACCCATTT 3841  
Db ATTATGATTTGAGGAAGAACATATCTGTGCTGAGGTGAGAAATGGCTTGGACCCATTT 4064  
3842 TGGGATATGATACAGACTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3901  
Db TGGGATATGATACAGACTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4122  
3902 CAAGGAATACAACTAG 3919  
Db CAAGGAATACAACTAG 4140

## RESULT 5

US-09-347-878-6  
; Sequence 6, Application US/09347878C  
; Patent No. 6376210

; GENERAL INFORMATION:  
; APPLICANT: Yuan, Chong  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES  
; FILE REFERENCE: 25885-1651  
; CURRENT APPLICATION NUMBER: US/09/347,878C  
; CURRENT FILING DATE: 1999-07-06  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 7224  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (395)..(4192)  
; FEATURES:  
; OTHER INFORMATION: Human methionine synthase  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: U73338/GenBank  
; US-09-347-878-6

Query Match 99.1%; Score 3883.2; DB 3; Length 7224;  
Best Local Similarity 99.8%; Pred.No. 0;  
Matches 3903; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 2 GTCACTCTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 61  
Db 333 GTCACTCTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 392  
Qy 62 ACATGTCAACCCGCGCTCCAGACCTGTGCAACCCGAGCTCTGAAGAAACCCCTGGGG 121  
Db 393 ACATGTCAACCCGCGCTCCAGACCTGTGCAACCCGAGCTCTGAAGAAACCCCTGGGG 452  
Qy 122 ATGAGATCAATGCTTCTGCAAGAGGATTTATGCTGCTGATGGGATGGGACCA 181  
Db 453 ATGAGATCAATGCTTCTGCAAGAGGATTTATGCTGCTGATGGGATGGGACCA 512



Qy 182 TGATCCAGCGGAGAGCTAAACGAAGAAACATCTCCAGAGTCAGGAATTTAAAGATCATG 241  
Db 513 TGATCCAGCGGAGAGCTAAACGAAGAAACATCTCCAGAGTCAGGAATTTAAAGATCATG 572  
Qy 242 CCAGGCGCTGAAAGGCAAAATAGCATTTTAAAGTATAACTCAGCCTGATGTCATTAAC 301  
Db 573 CCAGGCGCTGAAAGGCAAAATAGCATTTTAAAGTATAACTCAGCCTGATGTCATTAAC 632  
Qy 302 AAATCCATAAGGAATACCTGCTGGCTGGGCGAGATATCATTTGAAACAATACTTTAGCA 361  
Db 633 AAATCCATAAGGAATACCTGCTGGCTGGGCGAGATATCATTTGAAACAATACTTTAGCA 692  
Qy 362 GCATAGTATTTGCCAAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT 421  
Db 693 GCATAGTATTTGCCAAGCTGACTATGCGCTTTGAACACTTGGCCCTACCGGATGAACATGT 752  
Qy 422 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGTTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGTTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTGTTGGCAGGGCTCTGGGTCGGACTAATAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 813 GGTGTTGGCAGGGCTCTGGGTCGGACTAATAGACACTCTCTGTGTCCCATCTGTGG 872  
Qy 542 AAAGCGCGATTTATAGGAACATCAATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 601  
Db 873 AAAGCGCGATTTATAGGAACATCAATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 932  
Qy 602 CCAAAGGACTTCTGGATCGCGGGTTGATATCTTAATCAATGAAACTATTTTGTACTG 661  
Db 933 CCAAAGGACTTCTGGATCGCGGGTTGATATCTTAATCAATGAAACTATTTTGTACTG 992  
Qy 662 CCAATGCCAAGCAGCCTTGTGTCACCTCCAAATCTTTTGGAGAGAAATATGCTCCCC 721  
Db 993 CCAATGCCAAGCAGCCTTGTGTCACCTCCAAATCTTTTGGAGAGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTTACGGAGCATGCTGTTGATAAAGTGGCGGACTCTTTCCGGACAGA 781  
Db 1053 GGCCTATCTTTATTTACGGAGCATGCTGTTGATAAAGTGGCGGACTCTTTCCGGACAGA 1112  
Qy 782 CAGGAGAGGATTTGTCTACAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAAT 841  
Db 1113 CAGGAGAGGATTTGTCTACAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAAT 1172  
Qy 842 GTGCTTTGGTGCAGCTGAGATGAGACCTTTTATTTGAAATTAATGGAAAAATGTACACAG 901  
Db 1173 GTGCTTTGGTGCAGCTGAAATGAGACCTTTTATTTGAAATTAATGGAAAAATGTACACAG 1232  
Qy 902 CCTATGCTCTGTTATCCCAATGCAGGTCTTCCCAACACCTTTTGGTCACATGATGAAA 961  
Db 1233 CCTATGCTCTGTTATCCCAATGCAGGTCTTCCCAACACCTTTTGGTCACATGATGAAA 1292  
Qy 962 GGCCTTCTATGATGGCCACGACCTTAAAGGATTTTGTCTATGGATGGCTTGGTCAATATAG 1021  
Db 1293 GGCCTTCTATGATGGCCACGACCTTAAAGGATTTTGTCTATGGATGGCTTGGTCAATATAG 1352  
Qy 1022 TTGAGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAA 1081  
Db 1353 TTGAGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAA 1412  
Qy 1082 ATTTGAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTC 1141  
Db 1413 ATTTGAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTC 1472  
Qy 1142 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAATG 1201  
Db 1473 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAATG 1532  
Qy 1202 TTGAGAGGATCAAGGAAGTTTGTCTAACTCATCGCAGGAAACTATGAAGAGCCCTTGT 1261  
Db 1533 TTGAGAGGATCAAGGAAGTTTGTCTAACTCATCGCAGGAAACTATGAAGAGCCCTTGT 1592

Qy 1262 GTGTGTCCAAAGTGCAGGTGGAATATGGAGAGCCAGAGTGTGGATGTCAACATGATGATG 1321  
Db 1593 GTGTGTCCAAAGTGCAGGTGGAATATGGAGAGCCAGAGTGTGGATGTCAACATGATGATG 1652  
Qy 1322 GCATGCTAGATGTTCCAAAGTGCATATGACCAGATATTTTGCAACTTAATTTGCTTCCGAGCCAG 1381  
Db 1653 GCATGCTAGATGTTCCAAAGTGCATATGACCAGATATTTTGCAACTTAATTTGCTTCCGAGCCAG 1712  
Qy 1382 ACATCGCAAAAGGTACCTTTGTCATCGACTCTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1441  
Db 1713 ACATCGCAAAAGGTACCTTTGTCATCGACTCTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1772  
Qy 1442 TAAAGTCTCTCCAAAGGGAAGTCATTTGTCATATAGCATTAGTCTGAAGAAAGGAGAGACG 1501  
Db 1773 TAAAGTCTCTCCAAAGGGAAGTCATTTGTCATATAGCATTAGTCTGAAGAAAGGAGAGACG 1832  
Qy 1502 ACTTCTTTGGAGAGAGCGCCAGGAAGATTTAAAAAGTATGAGAGCTGCTATGGTGGTCAATGCGTT 1561  
Db 1833 ACTTCTTTGGAGAGAGCGCCAGGAAGATTTAAAAAGTATGAGAGCTGCTATGGTGGTCAATGCGTT 1892  
Qy 1562 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCT 1621  
Db 1893 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCT 1952  
Qy 1622 ACCATCTGCTTGTGAAAAAACTCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1681  
Db 1953 ACCATCTGCTTGTGAAAAAACTCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 2012  
Qy 1682 TCCTAACCATTTGGGACTGGGAATGGAGGAACACAACTTTGTATGCCATTAATTTTATCCATG 1741  
Db 2013 TCCTAACCATTTGGGACTGGGAATGGAGGAACACAACTTTGTATGCCATTAATTTTATCCATG 2072  
Qy 1742 CAACAAAGTCATTTAAAGAAACATTACTCGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 1801  
Db 2073 CAACAAAGTCATTTAAAGAAACATTACTCGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 2132  
Qy 1802 TGTCTTCTCTCTCCGAGGAATGGAGCCATTCGAGAGCAATGTCATGGGGTTTTCTCTTT 1861  
Db 2133 TGTCTTCTCTCTCCGAGGAATGGAGCCATTCGAGAGCAATGTCATGGGGTTTTCTCTTT 2192  
Qy 1862 ACCATGCAATCAAGTCTGGCATGGACATGAGAGATAGTGAATGCTGGAAAACTCCCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGGCATGGACATGAGAGATAGTGAATGCTGGAAAACTCCCTGTGT 2252  
Qy 1922 ATGATGATATCCATAAGGAACTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db 2253 ATGATGATATCCATAAGGAACTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
Qy 1982 CTGAGGCGCATGGAAGACTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGGAAGAAAG 2041  
Db 2313 CTGAGGCGCATGGAAGACTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGGAAGAAAG 2372  
Qy 2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTGAAGTATGCCCTTG 2101  
Db 2373 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTGAAGTATGCCCTTG 2432  
Qy 2102 TGAAGGCGCATGGAAGAAACATATTTGAGGATCTGAGGAGCCAGGTTAACCCAAAAA 2161  
Db 2433 TGAAGGCGCATGGAAGAAACATATTTGAGGATCTGAGGAGCCAGGTTAACCCAAAAA 2492  
Qy 2162 AATATCCCGGACCTCTCAATATATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2221  
Db 2493 AATATCCCGGACCTCTCAATATATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCCGGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCCGGGTTA 2612  
Qy 2282 TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACACACAG 2341  
Db 2613 TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACACACAG 2672  
Qy 2342 TGCTTAAAGGCGCAGTAGAGAGAGAGGAGCCCTTATCCAGGGGCCATCGTGTGCTGCCCATG 2401



|||||  
Db 2673 TGCCTTAACGGCAGTAGTAGAAGAGAGGCCCTTACAGGGGACCATCGTGTGCTGCCACTG 2732  
Qy 2402 TTAAAGGCGAGCTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGTGGCTGCAATA 2461  
Db 2733 TTAAAGGCGAGCTGCACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGTGGCTGCAATA 2792  
Qy 2462 ATTTCCGAGTTATTGATTAGGAGTCATGACTCATGTGATAGATAGTGAAGCTGCTC 2521  
Db 2793 ATTTCCGAGTTATTGATTAGGAGTCATGACTCATGTGATAGATAGTGAAGCTGCTC 2852  
Qy 2522 TTGACCAAAAGCAGATATAATGCGCTGTGAGAGCTCATGACTCATGCTTCCCTGATGAA 2581  
Db 2853 TTGACCAAAAGCAGATATAATGCGCTGTGAGAGCTCATGACTCATGCTTCCCTGATGAA 2912  
Qy 2582 TGATTTTGTGTTGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTTGTTGATGGAG 2641  
Db 2913 TGATTTTGTGTTGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATTTGTTGATGGAG 2972  
Qy 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCCACTG 2701  
Db 2973 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCCACTG 3032  
Qy 2702 TAATCCATGCTCTGACGCGTCCAGAGTGCTGTGTTCCAGCTGTTAGATGAA 2761  
Db 3033 TAATCCATGCTCTGACGCGTCCAGAGTGCTGTGTTCCAGCTGTTAGATGAA 3092  
Qy 2762 ATCTAAAGGATGAATACTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACC 2821  
Db 3093 ATCTAAAGGATGAATACTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACC 3152  
Qy 2822 ATTAGTGTCTCTCAAGGAGAGAGATACTTACCCCTTAAGTCAAGCCAGAGAAAAGTGGTT 2881  
Db 3153 ATTAGTGTCTCTCAAGGAGAGAGATACTTACCCCTTAAGTCAAGCCAGAGAAAAGTGGTT 3212  
Qy 2882 TCCAAATGGATTGGCTGTCTGAACCTCACAGTGAAGCCACCGCTTTATTTGGGACCCAGG 2941  
Db 3213 TCCAAATGGATTGGCTGTCTGAACCTCACAGTGAAGCCACCGCTTTATTTGGGACCCAGG 3272  
Qy 2942 TCTTTGAAGACTATGACCTCGCAGAGCTGTGGTACTACATTTGCTGGAAGCTTTCTTTTG 3001  
Db 3273 TCTTTGAAGACTATGACCTCGCAGAGCTGTGGTACTACATTTGCTGGAAGCTTTCTTTTG 3332  
Qy 3002 ATGTCTGCGAGCTCGGGGCAAGTACCGGAATCAGAGCTTCCCAAGATATTTAACGACA 3061  
Db 3333 ATGTCTGCGAGCTCGGGGCAAGTACCGGAATCAGAGCTTCCCAAGATATTTAACGACA 3392  
Qy 3062 AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3121  
Db 3393 AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3452  
Qy 3122 TGATTAGTCAAAAGAACTCCGGGCGCGGGGTGTGGTCTTGGCTGCGCAGCAGAGTA 3181  
Db 3453 TGATTAGTCAAAAGAACTCCGGGCGCGGGGTGTGGTCTTGGCTGCGCAGCAGAGTA 3512  
Qy 3182 TCCAGAGCAGATTTCACTGTACGAGAGGCTGTGTGCCCCAGGCTGCGAGGCCCATAG 3241  
Db 3513 TCCAGAGCAGATTTCACTGTACGAGAGGCTGTGTGCCCCAGGCTGCGAGGCCCATAG 3572  
Qy 3242 CCACCTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCCATCT 3301  
Db 3573 CCACCTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCCATCT 3632  
Qy 3302 ACTGCTCTCAGACTTCACTGCTCCCTTGCATCTTGGCATCCGTAACCTGCGGCTGT 3361  
Db 3633 ACTGCTCTCAGACTTCACTGCTCCCTTGCATCTTGGCATCCGTAACCTGCGGCTGT 3692  
Qy 3362 TTGCGGTGCTGCTTTGGGTAGAGAGCTGAGCAAGGCCCTATCAGGATGATGCTGACG 3421  
Db 3693 TTGCGGTGCTGCTTTGGGTAGAGAGCTGAGCAAGGCCCTATCAGGATGATGCTGACG 3752  
Qy 3422 ACTACAGCAGCATCATGTCTAAGGCGCTGGGGGACCGGCTGGCAGAGCCCTTTCCAGAG 3481  
|||||

Db 3753 ACTACAGCAGCATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTTCAGAG 3812  
Qy 3482 AGCTCCATGAAAGAGTTGCGCCGAGAACTGTGGGCTTACTGTGGAGTGAAGCTGAGCG 3541  
Db 3813 AGCTCCATGAAAGAGTTGCGCCGAGAACTGTGGGCTTACTGTGGAGTGAAGCTGAGCG 3872  
Qy 3542 TCCGAGAGCTGCGAAGGTTGCGGTACAAGGGCATCCGCCGGCTCCTGGCTACCCAGCC 3601  
Db 3873 TCCGAGAGCTGCGAAGGTTGCGGTACAAGGGCATCCGCCGGCTCCTGGCTACCCAGCC 3932  
Qy 3602 AGCCCGACCAACCCGAGAGCTCACCATTGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
Db 3933 AGCCCGACCAACCCGAGAGCTCACCATTGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3992  
Qy 3662 GCATTAGTTAAACAGAAATATTAGCAATGGCACTGCTTCAGCAGTCTCAGGCTCTACT 3721  
Db 3993 GCATTAGTTAAACAGAAATATTAGCAATGGCACTGCTTCAGCAGTCTCAGGCTCTACT 4052  
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGG 4112  
Qy 3782 ATTATGATTGAGCAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
Db 4113 ATTATGATTGAGCAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4172  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTCCTTTTATCTTGATGATCCT 3901  
Db 4173 TGGGATATGATACAGACTAAC--TTTTTTTTTTTTTTTTTTTTTTCCTTTTATCTTGATGATCCT 4230  
Qy 3902 CAAGGAAATACACCTAG 3919  
Db 4231 CAAGGAAATACACCTAG 4248

## RESULT 6

US-09-962-665-1

; Sequence 1, Application US/09962665

; Patent No. 6537759

; GENERAL INFORMATION:

; APPLICANT: Stanton, Jr., Vincent P.

; TITLE OF INVENTION: POLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE

; TITLE OF INVENTION: TREATMENT OF DISEASE

; FILE REFERENCE: 11926-015004

; CURRENT APPLICATION NUMBER: US/09/962,665

; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 09/658,659

; PRIOR FILING DATE: 2000-09-08

; PRIOR APPLICATION NUMBER: 09/596,033

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 09/357,743

; PRIOR FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: 09/357,024

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: 60/093,484

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 7224

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 194..3209

; OTHER INFORMATION: n = c or g

; NAME/KEY: misc feature

; LOCATION: 1136..1334, 3150, 5551, 5934

; OTHER INFORMATION: n = a or g

; NAME/KEY: misc feature

; LOCATION: 284, 1252, 1699, 5573, 5659, 5678, 5874

; OTHER INFORMATION: n = c or t





```

; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 194..3209
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1136..1334, 3150, 5551, 5934
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 284, 1252, 1699, 5573, 5659, 5678, 5874
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 3207
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 5444
; OTHER INFORMATION: n = c or a
US-09-963-333-1

Query Match      98.9%; Score 3876.2; DB 3; Length 7224;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3896; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

QY 2  GTCACTCTGTGAGAGCAGCTTCTCTCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61
DB 333  GTCACTCTGTGAGAGCAGCTTCTCTCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392

QY 62  ACATGTCAACCGCGCTCCAAAGCTGTTCGCAACCGCAAGGCTCGAAGAAACCTCGCGG 121
DB 393  ACATGTCAACCGCGCTCCAAAGCTGTTCGCAACCGCAAGGCTCGAAGAAACCTCGCGG 452

QY 122  ATGAGATCAATGCCATCTTCGAGAGAGGATTTATGGCTCGATGGAGGGATGGGACCA 181
DB 453  ATGAGATCAATGCCATCTTCGAGAGAGGATTTATGGCTCGATGGAGGGATGGGACCA 512

QY 182  TGATCCAGCGGAGAGCTTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241
DB 513  TGATCCAGCGGAGAGCTTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 572

QY 242  CCAGCGCGCTGAAGCAACATGACATTTTAAAGTATTAATCACTCAGCTGATGTCAATTACC 301
DB 573  CCAGCGCGCTGAAGCAACATGACATTTTAAAGTATTAATCACTCAGCTGATGTCAATTACC 632

QY 302  AAATCCATAGGAATACCTGTGGCTGGGCGAGATATCATTTGAACAAATACCTTTAGCA 361
DB 633  AAATCCATAGGAATACCTGTGGCTGGGCGAGATATCATTTGAACAAATACCTTTAGCA 692

QY 362  GCACCTAGTATTGCCCAAGCTGACTATGCTTGAACACTTGGCCCTACCGGATGAACATGT 421
DB 693  GCACCTAGTATTGCCCAAGCTGACTATGCTTGAACACTTGGCCCTACCGGATGAACATGT 752

QY 422  GCTCTGAGGAGTGCCGAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGAAATTAAGA 481
DB 753  GCTCTGAGGAGTGCCGAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGAAATTAAGA 812

QY 482  GGTTTGTGGCAGGGCTCTGGGTCCGACTATTAAGACACTCTCTGTGTCCTCCATCTGTGG 541
DB 813  GGTTTGTGGCAGGGCTCTGGGTCCGACTATTAAGACACTCTCTGTGTCCTCCATCTGTGG 872

QY 542  AAAGCCCGATTTATAGGAACATCACTTTGTAGAGCTTTGTGAAGCATACCAAGAGCAGG 601
DB 943  AAAGCCCGATTTATAGGAACATCACTTTGTAGAGCTTTGTGAAGCATACCAAGAGCAGG 932

602  CCAAGGACCTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGATCTG 661
933  CCAAGGACCTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGATCTG 992
662  CCAATGCAAGGAGCAGCTTGTGTCACCTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721
993  CCAATGCAAGGAGCAGCTTGTGTCACCTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052
722  GGCTATCTTTATTTTCAAGGACGATCGTTGATTTAAAGTGGCGGACCTCTTCCGACAGA 781
1053  GGCTATCTTTATTTTCAAGGACGATCGTTGATTTAAAGTGGCGGACCTCTTCCGACAGA 1112
782  CAGGAGGAGGATTTGTCATCAGCGGTCTCATGAGAACCACTCTGCTCATTTGGATTAAT 841
1113  CAGGAGGAGGATTTGTCATCAGCGGTCTCATGAGAACCACTCTGCTCATTTGGATTAAT 1172
842  GTGCTTTGGGTGACGCTGAGATGAGACCTTTTATTTGAAATAAATTTGGAATAATGTACAACAG 901
1173  GTGCTTTGGGTGACGCTGAGATGAGACCTTTTATTTGAAATAAATTTGGAATAATGTACAACAG 1232
902  CCTATGTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTGGTGTGACTATGATGAAA 961
1233  CCTATGTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTGGTGTGACTATGATGAAA 1292
962  CGCCTTCTATGATGGCGCAAGCAGCTTAAAGGATTTTGTATGATGGCTTGGTCAATATAG 1021
1293  CGCCTTCTATGATGGCGCAAGCAGCTTAAAGGATTTTGTATGATGGCTTGGTCAATATAG 1352
1022  TTGAGGATGCTGTGGGTCAACACCGAGATCATATCAAGGAAATTTGCTGAAGAGCTGTGAAA 1081
1353  TTGAGGATGCTGTGGGTCAACACCGAGATCATATCAAGGAAATTTGCTGAAGAGCTGTGAAA 1412
1082  ATTGTAAAGCTAGAGTTCCACCTGCGACCTGCTTTTGAAGACATATGTTTACTGTCTGGTC 1141
1413  ATTGTAAAGCTAGAGTTCCACCTGCGACCTGCTTTTGAAGACATATGTTTACTGTCTGGTC 1472
1142  TAGAGCCCTTCAGGATTTGGACCTGTACACCACTTTGTTTAACATTTGGAGGCGCTGTAATG 1201
1473  TAGAGCCCTTCAGGATTTGGACCTGTACACCACTTTGTTTAACATTTGGAGGCGCTGTAATG 1532
1202  TTGCAGGATCAAGGAAAGTTTGTCTAAACTCATCATGCGAGGAAACTATGAAGAACCTTGT 1261
1533  TTGCAGGATCAAGGAAAGTTTGTCTAAACTCATCATGCGAGGAAACTATGAAGAACCTTGT 1592
1262  GTGTTCGCAAGTGCAGGTGGAAATGGAGCCAGGCTTGGATGTGCAATGATGATG 1321
1593  GTGTTCGCAAGTGCAGGTGGAAATGGAGCCAGGCTTGGATGTGCAATGATGATG 1652
1322  GCATGCTAGATGGTCCAAAGTGCATGACCAAGTTTTCGAACCTTAAATGCTTCCAGGCGAG 1381
1653  GCATGCTAGATGGTCCAAAGTGCATGACCAAGTTTTCGAACCTTAAATGCTTCCAGGCGAG 1712
1382  ACATCGCAAGGATACCTTTGTGCACTCGACTCTTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1441
1713  ACATCGCAAGGATACCTTTGTGCACTCGACTCTTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1772
1442  TAAAGTCTGCGCAAGGAAAGTGCATTTGTCATAGCATTAGCTTGAAGGAGGAGGAGACG 1501
1773  TAAAGTCTGCGCAAGGAAAGTGCATTTGTCATAGCATTAGCTTGAAGGAGGAGGAGACG 1832
1502  ACTTCTTGGAGAGGCGCAGGAGATTTAAAGATGATGAGCTGCTATGGTGGTCAATGGCTT 1561
1833  ACTTCTTGGAGAGGCGCAGGAGATTTAAAGATGATGAGCTGCTATGGTGGTCAATGGCTT 1892
1562  TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCAACCGGGGCT 1621
1893  TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATAATCAGAGTGTGCAACCGGGGCT 1952
1622  ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1681
1953  ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 2012
```

Qy	1682	TCCTAAACATTTGGGACTGGATGAGGAAACACAACTTTGTATGCCATTAATTTTATCCATG	1741
Dd	2013	TCCTAAACATTTGGGACTGGATGAGGAAACAACTTTGTATGCCATTAATTTTATCCATG	2072
Qy	1742	CAACAAAAGTCATTTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGGAGGCTCTTTCCAACT	1801
Dd	2073	CAACAAAAGTCATTTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGGAGGCTCTTTCCAACT	2132
Qy	1802	TGTCCTCTCTCTCCGAGGAATGAAGCAATTCGAGAAAGCAATGCATGGGGTTTTCCTTT	1861
Dd	2133	TGTCCTCTCTCTCCGAGGAATGAAGCAATTCGAGAAAGCAATGCATGGGGTTTTCCTTT	2192
Qy	1862	ACCATGCAATCAAGTCTGGCATGACATGAGATAGTGAATGCTGGAAACCTCCTCTGT	1921
Dd	2193	ACCATGCAATCAAGTCTGGCATGACATGAGATAGTGAATGCTGGAAACCTCCTCTGT	2252
Qy	1922	ATGATGATATCCATAAGGAACCTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC	1981
Dd	2253	ATGATGATATCCATAAGGAACCTCTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC	2312
Qy	1982	CTGAGGCCACTGAGAAGCTTTACGTTATGCCAGACTCAAGGCACAGGAGGGAGAAAG	2041
Dd	2313	CTGAGGCCACTGAGAAGCTTTACGTTATGCCAGACTCAAGGCACAGGAGGGAGAAAG	2372
Qy	2042	TCATTGACATGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCTTG	2101
Dd	2373	TCATTGACATGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCTTG	2432
Qy	2102	TGAAGGCAATTTGAAAAACATTAATTGAGGATACCTGAGGAAGCCAGGTTAAACCAAAAA	2161
Dd	2433	TGAAGGCAATTTGAAAAACATTAATTGAGGATACCTGAGGAAGCCAGGTTAAACCAAAAA	2492
Qy	2162	AATATCCCGACCTCTCAATATAATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2221
Dd	2493	AATATCCCGACCTCTCAATATAATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2552
Qy	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGGTTA	2281
Dd	2553	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGGTTA	2612
Qy	2282	TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAAAGCAAGAG	2341
Dd	2613	TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAAAGCAAGAG	2672
Qy	2342	TGCTTTAACGGCAAGTAGAAGAGAGGACCCCTTACAGGGCACCACATCGTGTGGCCACTG	2401
Dd	2673	TGCTTTAACGGCAAGTAGAAGAGAGGACCCCTTACAGGGCACCACATCGTGTGGCCACTG	2732
Qy	2402	TTAAAGGCGACGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA	2461
Dd	2733	TTAAAGGCGACGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA	2792
Qy	2462	ATTTCCGAGTTATTGATTAGGAGTCACTCATGTGATAGATACTGAAAGCTGCTC	2521
Dd	2793	ATTTCCGAGTTATTGATTAGGAGTCACTCATGTGATAGATACTGAAAGCTGCTC	2852
Qy	2522	TTGACCAACAAAGCAGATATAATTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2581
Dd	2853	TTGACCAACAAAGCAGATATAATTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2912
Qy	2582	TGATTTTGTGTTGCCAAGGAAATGGAGATAGTCTATAAGGATTTCCATTGTTGATGGAG	2641
Dd	2913	TGATTTTGTGTTGCCAAGGAAATGGAGATAGTCTATAAGGATTTCCATTGTTGATGGAG	2972
Qy	2642	GAGCAACCACTTCAAAAACCCACACAGCAGTAAATAAGCTCCGAGATACAGTGACCTG	2701
Dd	2973	GAGCAACCACTTCAAAAACCCACACAGCAGTAAATAAGCTCCGAGATACAGTGACCTG	3032
Qy	2702	TAATCCATGCTCTGAGCGCTCCAAAGAGTGGTGTGTTTCCAGCTGTTAGATGAAA	2761
Dd	3033	TAATCCATGCTCTGAGCGCTCCAAAGAGTGGTGTGTTTCCAGCTGTTAGATGAAA	3092

Qy	2762	ATCTAAAGGATGAAATACCTTTGAGGAATCATGGAATATGAGATATTTAGACGAGCC	2821
Dd	3093	ATCTAAAGGATGAAATACCTTTGAGGAATCATGGAATATGAGATATTTAGACGAGCC	3152
Qy	2822	ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAACTCAAGCCAGAAAAAGTGTT	2881
Dd	3153	ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAACTCAAGCCAGAAAAAGTGTT	3212
Qy	2882	TCCAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACAGTTTATTTGGGACCCAGG	2941
Dd	3213	TCCAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACAGTTTATTTGGGACCCAGG	3272
Qy	2942	TCCTTTGAAGACTATGACCTGCAAGAGCTGTGAGCTACATTTGACTGGAAGCCCTTTCTTG	3001
Dd	3273	TCCTTTGAAGACTATGACCTGCAAGAGCTGTGAGCTACATTTGACTGGAAGCCCTTTCTTG	3332
Qy	3002	ATGCTGGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTTCCCAAGATATTTAAACGACA	3061
Dd	3333	ATGCTGGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTTCCCAAGATATTTAAACGACA	3392
Qy	3062	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCACATATGCTGAACACAC	3121
Dd	3393	AAACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCACATATGCTGAACACAC	3452
Qy	3122	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGTCTGGTTGCTGGCCAGCAAGAGTA	3181
Dd	3453	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGTCTGGTTGCTGGCCAGCAAGAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACGAGAGGCTGCTGTGCCCCAGGCTGCAGAGCCCATAG	3241
Dd	3513	TCCAAGACGACATTCACCTGTACGCGGAGGCTGCTGTGCCCCAGGCTGCAGAGCCCATAG	3572
Qy	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCACGGAGCCATACT	3301
Dd	3573	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCACGGAGCCATACT	3632
Qy	3302	ACTGCCCTCTCAGACTTCATCGCTCCCTTGGCATTTCTGGCATCCGTGACTACTGGGCCCTGT	3361
Dd	3633	ACTGCCCTCTCAGACTTCATCGCTCCCTTGGCATTTCTGGCATCCGTGACTACTGGGCCCTGT	3692
Qy	3362	TTGCCGTTGCTCTCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG	3421
Dd	3693	TTGCCGTTGCTCTCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG	3752
Qy	3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGACAGAG	3481
Dd	3753	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGACAGAG	3812
Qy	3482	AGCTCCATGAAAGAGTTCCGCCGAGAACTGTGGGCTCTACTGTGGCAGTGAGCAGCTGACG	3541
Dd	3813	AGCTCCATGAAAGAGTTCCGCCGAGAACTGTGGGCTCTACTGTGGCAGTGAGCAGCTGACG	3872
Qy	3542	TCGAGAGCTGCCGAGGTTGCGGCTCAAGGGGACCGGCTGGCAGAGGCGCTTTGACAGAG	3601
Dd	3873	TCGAGAGCTGCCGAGGTTGCGGCTCAAGGGGACCGGCTGGCAGAGGCGCTTTGACAGAG	3932
Qy	3602	AGCCCCACCAACCGGAGAGCTCACCATGTGGAGACTTCGAGAGCATCGAGCAGTCTACAG	3661
Dd	3933	AGCCCCACCAACCGGAGAGCTCACCATGTGGAGACTTCGAGAGCATCGAGCAGTCTACAG	3992
Qy	3662	GCATTTAGGTTAAACAGATCATTTAGCAATGGCACTGCTGTCAGCAAGTCTCAGGCTCTACT	3721
Dd	3993	GCATTTAGGTTAAACAGATCATTTAGCAATGGCACTGCTGTCAGCAAGTCTCAGGCTCTACT	4052
Qy	3722	TCCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAGGATCAGGTTGAGG	3781
Dd	4053	TCCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTCCAGGATCAGGTTGAGG	4112
Qy	3782	ATTATGATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTTGGACCCATTT	3841
Dd	4113	ATTATGATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTTGGACCCATTT	4172
Qy	3842	TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTATCTTGATGATCCT	3901



Db 4173 TGGGATATGATACAGACTAAC--TTTTTTTTTTTTTGGCTTTTTTATTTCTTGATGATCCT 4230  
Qy 3902 CAAGGAATACACCTAG 3919  
Db 4231 CAAGGAATACACCTAG 4248  
RESULT 8  
US-09-962-677-1  
; Sequence 1, Application US/09962677  
; Patent No. 6759200  
; GENERAL INFORMATION:  
; APPLICANT: Stanton, Jr., Vincent P.  
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE  
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING  
; TITLE OF INVENTION: THE TREATMENT OF DISEASE  
; FILE REFERENCE: 11926-015003  
; CURRENT APPLICATION NUMBER: US/09/962,677  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/658,659  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 09/596,033  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 09/357,743  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 09/357,024  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: 60/093,484  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 7224  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 194..3209  
; OTHER INFORMATION: n = c or g  
; NAME/KEY: misc feature  
; LOCATION: 1136, 1334, 3150, 5551, 5934  
; OTHER INFORMATION: n = a or g  
; NAME/KEY: misc feature  
; LOCATION: 284..1252, 1699, 5573, 5659, 5678, 5874  
; OTHER INFORMATION: n = c or t  
; NAME/KEY: misc feature  
; LOCATION: 3207  
; OTHER INFORMATION: n = g or t  
; NAME/KEY: misc feature  
; LOCATION: 5444  
; OTHER INFORMATION: n = c or a  
US-09-962-677-1  
Query Match 98.9%; Score 3876.2; DB 3; Length 7224;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3896; Conservative 0; Mismatches 20; Indels 2; Gaps 1;  
Qy 2 GTCACCTGTGGAGAGACGCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61  
Db 333 GTCACCTGTGGAGAGACGCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392  
Qy 62 ACATGTCACCGCGCTCCAAAGACCTGTCGCAACCCGAAAGGTCGAAAGAAACCCCTGCGGG 121  
Db 393 ACATGTCACCGCGCTCCAAAGACCTGTCGCAACCCGAAAGGTCGAAAGAAACCCCTGCGGG 452  
Qy 122 ATGAGATCAATGCCATTCTGCAGAGAGAGATTATGCTGCTGGATGGAGGATGGGGACCA 181  
Db 453 ATGAGATCAATGCCATTCTGCAGAGAGAGATTATGCTGCTGGATGGAGGATGGGGACCA 512  
Qy 182 TGATCCAGGGGAGAGCTAAACGAGAACACTCCGAGGTCAGGAATTTAAAGATCATG 241  
Db 513 TGATCCAGGGGAGAGCTAAACGAGAACACTCCGAGGTCAGGAATTTAAAGATCATG 572

Qy 242 CCAGGCGCTGAAAGGCAACATGACATTTTAAAGTATACTCAGCCTGATGTCTATTACC 301  
Db 573 CCAGGCGCTGAAAGGCAACATGACATTTTAAAGTATACTCAGCCTGATGTCTATTACC 632  
Qy 302 AAATCCATAAGGAATACTTCTGCTGGGCGAGANATCATTTGAAAACAAATATCTTTTAGCA 361  
Db 633 AAATCCATAAGGAATACTTCTGCTGGGCGAGATATCATTTGAAAACAAATATCTTTTAGCA 692  
Qy 362 GCCTAGTATTGCCCAAGCTGACTATGGCCCTTGAACACTTTGGCCCTACCGATGAACATGT 421  
Db 693 GCCTAGTATTGCCCAAGCTGACTATGGCCCTTGAACACTTTGGCCCTACCGATGAACATGT 752  
Qy 422 GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGAGAGGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGAGAGGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 813 GGTTTGTGGCAGGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG 872  
Qy 542 AAAGGCCGGATTATAGGAACATCACATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 601  
Db 873 AAAGGCCGGATTATAGGAACATCACATTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGG 932  
Qy 602 CCAAAGGACTTCTGGATGGCGGGTGTGATATCTTTACTCATTTGAAACTATTTTGTGATCTG 661  
Db 932 CCAAAGGACTTCTGGATGGCGGGTGTGATATCTTTACTCATTTGAAACTATTTTGTGATCTG 992  
Qy 662 CCAATGCAAGGCGCCTTGTGTGACTCCAAATCTTTTGAAGGAGAAATATGCTCCCC 721  
Db 993 CCAATGCAAGGCGCCTTGTGTGACTCCAAATCTTTTGAAGGAGAAATATGCTCCCC 1052  
Qy 722 GGCTATCTTTATTTCAGGAGCGATCGTTGATATAAGTGGCGGACCTTTCCGAGCAGA 781  
Db 1053 GGCTATCTTTATTTCAGGAGCGATCGTTGATATAAGTGGCGGACCTTTTCGAGCAGA 1112  
Qy 782 CAGGAGAGCGATTGTCTCATCAGCGTGTCTCATGAGAACCACTCTGCTGATTTGATTAAT 841  
Db 1113 CAGGAGAGCGATTGTCTCATCAGCGTGTCTCATGAGAACCACTCTGCTGATTTGATTAAT 1172  
Qy 842 GTGCTTTGGGTGCGAGCTGAGATGAGACCTTTTATTTGAAATAATTTGGAATAATGTAACAG 901  
Db 1173 GTGCTTTGGGTGCGAGCTGAAATGAGACCTTTTATTTGAAATAATTTGGAATAATGTAACAG 1232  
Qy 902 CCTATGTCCTGTGTTATCCCAATGCGAGTCTTCCCAACACTTTGGTGACATATGATGAAA 961  
Db 1233 CCTATGTCCTGTGTTATCCCAATGCGAGTCTTCCCAACACTTTGGTGACATATGATGAAA 1292  
Qy 962 CGCCTTCTATGATGCGCAAGCACCTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
Db 1293 CGCCTTCTATGATGCGCAAGCACCTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1352  
Qy 1022 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTGTAAGCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGTC 1141  
Db 1413 ATTGTAAGCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGTC 1472  
Qy 1142 TAGAGCCCTTCAGGATTTGGACCGGTACCAACTTTTGTAAATTTGGAGAGCGCTGTAAATG 1201  
Db 1473 TAGAGCCCTTCAGGATTTGGACCGGTACCAACTTTTGTAAATTTGGAGAGCGCTGTAAATG 1532  
Qy 1202 TTGAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAAACTATGAAAGAGCCTTGT 1261  
Db 1533 TTGAGGATCAAGGAAGTTTGTCTAAACTCATATGCGAGGAAACTATGAAAGAGCCTTGT 1592  
Qy 1262 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCCGAGGTTGGATGTCACCATGATGATG 1321  
Db 1593 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCCGAGGTTGGATGTCACCATGATGATG 1652



Qy	1322	GCATGCTAGATGGTCCAAAGTGCAGATGACAGANTTTTGCAACTTAAATTTGCTTCCGAGCCAG	1381
Db	1653	GCATGCTAGATGGTCCAAAGTGCAGATGACAGANTTTTGCAACTTAAATTTGCTTCCGAGCCAG	1712
Qy	1382	ACATCGCAAGAGTACCTTTTGCGCATCGACTCCTCCAAATTTTGTGCTGCTGANTTGAAGCTGGGT	1441
Db	1713	ACATCGCAAGAGTACCTTTTGCGCATCGACTCCTCCAAATTTTGTGCTGCTGANTTGAAGCTGGGT	1772
Qy	1442	TAAAGTGTCTGCAAGGGGAAGTGCATTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGGACG	1501
Db	1773	TAAAGTGTCTGCAAGGGGAAGTGCATTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGGACG	1832
Qy	1502	ACTTCTTGGAGAGCCACGGAAGATATTAAGATATGAGCTGCTATGCTGGTCTATGCTGT	1561
Db	1833	ACTTCTTGGAGAGCCACGGAAGATATTAAGATATGAGCTGCTATGCTGGTCTATGCTGT	1892
Qy	1562	TTGATGAAGAGGACAGGACACAGAACAGACACAAAATCAGAGTGTGCACCCGGGCGCT	1621
Db	1893	TTGATGAAGAGGACAGGACACAGAACAGACACAAAATCAGAGTGTGCACCCGGGCGCT	1952
Qy	1622	ACCATCTGCTTTGTAAGAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCCCTAATA	1681
Db	1953	ACCATCTGCTTTGTAAGAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCCCTAATA	2012
Qy	1682	TCCTAAACATTTGGGACTGGAAATGAGAGAACACAACCTTGTATGTCATTAATTTTATCCATG	1741
Db	2013	TCCTAAACATTTGGGACTGGAAATGAGAGAACACAACCTTGTATGTCATTAATTTTATCCATG	2072
Qy	1742	CAACAAAGTCATTAAAGAACATTACTGTGAGCCAGAAATAGTGGAGGTCTTTCCCAACT	1801
Db	2073	CAACAAAGTCATTAAAGAACATTACTGTGAGCCAGAAATAGTGGAGGTCTTTCCCAACT	2132
Qy	1802	TGTCTCTTCTCTCCGAGGAATGGAAGCCATTTCGAGAGCAATTCATCGGGGTCTTCCCTTT	1861
Db	2133	TGTCTCTTCTCTCCGAGGAATGGAAGCCATTTCGAGAGCAATTCATCGGGGTCTTCCCTTT	2192
Qy	1862	ACCATGCAATCAAGTCTGGCATGGAATGAGAGATAGTGAATGCTGGGAAACCTCCCTGTGT	1921
Db	2193	ACCATGCAATCAAGTCTGGCATGGAATGAGAGATAGTGAATGCTGGGAAACCTCCCTGTGT	2252
Qy	1922	ATGATGATATCCATAAGGAACCTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACC	1981
Db	2253	ATGATGATATCCATAAGGAACCTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACC	2312
Qy	1982	CTGAGGCCACTGAGAAGCTCTTAGTTATGCCCAGACTCAAGGCACAGGAGGGGAAGAAG	2041
Db	2313	CTGAGGCCACTGAGAAGCTCTTAGTTATGCCCAGACTCAAGGCACAGGAGGGGAAGAAG	2372
Qy	2042	TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCCCTTGAGTATGCCCTTG	2101
Db	2373	TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCCCTTGAGTATGCCCTTG	2432
Qy	2102	TGAAGGGCATTTGAAACACATATTATTGAGGATCTGTGAAGAGCCAGTTTAAACCAAAAAA	2161
Db	2433	TGAAGGGCATTTGAAACACATATTATTGAGGATCTGTGAAGAGCCAGTTTAAACCAAAAAA	2492
Qy	2162	AATATCCCGGACCTCTCAATATAATGAAGAGCCCTGTGATGAATGGAATGAAATTCGTG	2221
Db	2493	AATATCCCGGACCTCTCAATATAATGAAGAGCCCTGTGATGAATGGAATGAAATTCGTG	2552
Qy	2222	GTGATCTTTTTGGAGCTGGAAAAATGTTTTACCTCAGGTTTATAAGTCCAGCCCGGGTTA	2281
Db	2553	GTGATCTTTTTGGAGCTGGAAAAATGTTTTACCTCAGGTTTATAAGTCCAGCCCGGGTTA	2612
Qy	2282	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAAACCAAGAG	2341
Db	2613	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAAACCAAGAG	2672
Qy	2342	TGCTTTAACGGCACAGTAGAAGAGGAGCCCTTACCAGGGCACCCTCGTCTGCGCCACTG	2401
Db	2673	TGCTTTAACGGCACAGTAGAAGAGGAGCCCTTACCAGGGCACCCTCGTCTGCGCCACTG	2732
Qy	2402	TTAAAGGCGACGTGCACGACATAGGCAAGAACATAGTTTGAAGTAGTCTTGGCTGCAATA	2461

Db	2733	TTAAAGCGCAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGGCAATA	2792
Qy	2462	ATTTCGCGAGTTATTGATTTAGGAGTCATGACTCCATGTGATAAGATACTGAAAGCTGCTC	2521
Db	2793	ATTTCCGAGTTATTGATTTAGGAGTCATGACTCCATGTGATAAGATACTGAAAGCTGCTC	2852
Qy	2522	TTGACCAAAAGCAGATATAATTGGCTGTGAGGACTCATCTCTTCCTCGATGAAA	2581
Db	2853	TTGACCAAAAGCAGATATAATTGGCTGTGAGGACTCATCTCTTCCTCGATGAAA	2912
Qy	2582	TGATTTTGTGTGCCAAGGAATGGAGAGATTAGCTATAGGATTCATTGTTGATTTGGAG	2641
Db	2913	TGATTTTGTGTGCCAAGGAATGGAGAGATTAGCTATAGGATTCATTGTTGATTTGGAG	2972
Qy	2642	GAGCAACCACTTCAAAAACCAACACAGCAGTATAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2973	GAGCAACCACTTCAAAAACCAACACAGCAGTATAAATAGCTCCGAGATACAGTGCACCTG	3032
Qy	2702	TAATPCCATGCTCTGGACCGTCCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA	2761
Db	3033	TAATPCCATGCTCTGGACCGTCCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA	3092
Qy	2762	ATCTAAAGGATGAATACTTTTGAGGAATCATGGAAGATATGAAGATATTAGACAGGACC	2821
Db	3093	ATCTAAAGGATGAATACTTTTGAGGAATCATGGAAGATATGAAGATATTAGACAGGACC	3152
Qy	2822	ATTATGAGTCTCTCAAGGAGAGGATACATTACCTTTAGCTCAAGCCAGAAAAAGTGGTT	2881
Db	3153	ATTATGAGTCTCTCAAGGAGAGGATACATTACCTTTAGCTCAAGCCAGAAAAANTGTT	3212
Qy	2882	TCCAAATGGATTTGGCTGCTGTAACCTCACCCAGTGAAGCCACGTTTATTGGGACCCAGG	2941
Db	3213	TCCAAATGGATTTGGCTGCTGTAACCTCACCCAGTGAAGCCACGTTTATTGGGACCCAGG	3272
Qy	2942	TCCTTTGAAGACTATGACTCTGCAGAAAGCTGTGTGACTACATTGACTGGAAGCCTTTCTTTG	3001
Db	3273	TCCTTTGAAGACTATGACTCTGCAGAAAGCTGTGTGACTACATTGACTGGAAGCCTTTCTTTG	3332
Qy	3002	ATGTCTCGCAGCTCCGGGGCAAGTACCAGGAATCGAGGCTTCCCCAAGATATTTAAACGACA	3061
Db	3333	ATGTCTCGCAGCTCCGGGGCAAGTACCAGGAATCGAGGCTTCCCCAAGATATTTAAACGACA	3392
Qy	3062	AAAACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCCAACATATCTCTGAAACAC	3121
Db	3393	AAAACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCCCAACATATCTCTGAAACAC	3452
Qy	3122	TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGGTTCTGGCCAGCACAGAGTA	3181
Db	3453	TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGGTTCTGGCCAGCACAGAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACGAGAGGCTCTGTGCCCCAGGCTGCAGAGGCCCATAG	3241
Db	3513	TCCAAGACGACATTCACCTGTACGAGAGGCTCTGTGCCCCAGGCTGCAGAGGCCCATAG	3572
Qy	3242	CCACTTTCTATGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAACGAGGCCATACT	3301
Db	3573	CCACTTTCTATGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAACGAGGCCATACT	3632
Qy	3302	ACTGCCTCTCAGACTTCATCGCTCCCTTTGCATTCTGGCATPCGTGTACTACCTGGGCTGT	3361
Db	3633	ACTGCCTCTCAGACTTCATCGCTCCCTTTGCATTCTGGCATPCGTGTACTACCTGGGCTGT	3692
Qy	3362	TTGCCGTTGCCCTTTGGGGTAGAGAGCTGAGCAAGGCTATAGGATGATGGTGACG	3421
Db	3693	TTGCCGTTGCCCTTTGGGGTAGAGAGCTGAGCAAGGCTATAGGATGATGGTGACG	3752
Qy	3422	ACTACACGACATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCGAGAG	3481
Db	3753	ACTACACGACATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCGAGAG	3812
Qy	3482	AGCTCCATGAAAGATTCCGCGAGAACTGTGGGCCCTACTGTGGCAGTGACAGCTGGAAG	3541

Db 3813 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCGCTACTGTGGCAGTGAGCAGCTGGAGC 3872  
Qy  
3542 TCGCAGACCTCGGAGGTTGGGTACAGGGCATCCGCGCGCTCCGGCTACCCAGCC 3601  
Db  
3873 TCGCAGACCTCGGAGGTTGGGTACAGGGCATCCGCGCGCTCCGGCTACCCAGCC 3932  
Qy  
3602 AGCCCGACACACCGAGAGCTCACCAATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
Db  
3933 AGCCCGACACACCGAGAGCTCACCAATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3992  
Qy  
3662 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGCTGTTCAGCAGTCTCAGCCCTCTACT 3721  
Db  
3993 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGCTGTTCAGCAGTCTCAGCCCTCTACT 4052  
Qy  
3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAGATTTCCAGAGATCAGGTTGAGG 3781  
Db  
4053 TCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAGATTTCCAGAGATCAGGTTGAGG 4112  
Qy  
3782 ATTATGCAATCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
Db  
4113 ATTATGCAATCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4172  
Qy  
3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTGTGCTTTTTTATCTTTGATGATCCT 3901  
Db  
4173 TGGGATATGATACAGACTAAC - TTTTTTTTTTTTTTTTTTGTGCTTTTTTATCTTTGATGATCCT 4230  
Qy  
3902 CAAGGAATACACCTAG 3919  
Db  
4231 CAAGGAATACACCTAG 4248

RESULT 9

US-08-980-326-75  
; Sequence 75, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 75  
; LENGTH: 3856  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: (1)...(1265)  
; OTHER INFORMATION: Coding sequence for mutant methionine synthases:  
; OTHER INFORMATION: 2640-2642 can be AAT or deleted; 2756 can be A or  
; OTHER INFORMATION: G:2758 can be C or G.  
US-08-980-326-75

Query Match 98.3%; Score 3851; DB 3; Length 3856;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3851; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 64 ATGTCAACCGCGCTCCAAGACTGTGCGAACCCCGAAGTCTGGAAGAAACCCCTGGGGAT 123  
Db 1 ATGTCAACCGCGCTCCAAGACTGTGCGAACCCCGAAGTCTGGAAGAAACCCCTGGGGAT 60

Qy 124 GAGATCAATGCCATTTCTGCAGAGAGGATTTATGTTGCTGATGAGGATGGGACCATG 183  
Db  
61 GAGATCAATGCCATTTCTGCAGAGAGGATTTATGTTGCTGATGAGGATGGGACCATG 120  
Qy 184 ATCCAGCGGAGAGCTTAAACGAGAACACATTCGAGGTCAGGAATTTAAAGATCATGCC 243  
Db  
121 ATCCAGCGGAGAGCTTAAACGAGAACACATTCGAGGTCAGGAATTTAAAGATCATGCC 180  
Qy 244 AGCGCGTGAAGGCAACATGACATTTTAAGTATTAATCTCAGCTGATGTCATTTCACAA 303  
Db  
181 AGCGCGTGAAGGCAACATGACATTTTAAGTATTAATCTCAGCTGATGTCATTTCACAA 240  
Qy 304 ATCCATAAGGAATACTTGTCTGGCTGGGCGAGATATCAATGAAAACAAATACTTTTAGCAGC 363  
Db  
241 ATCCATAAGGAATACTTGTCTGGCTGGGCGAGATATCAATGAAAACAAATACTTTTAGCAGC 300  
Qy 364 ACTAGTATTCGCCAAGCTGACTATGCGCTTTGAAACATCTTGCGCTTACCGGATGAACATGTC 423  
Db  
301 ACTAGTATTCGCCAAGCTGACTATGCGCTTTGAAACATCTTGCGCTTACCGGATGAACATGTC 360  
Qy 424 TCTGCAGAGTGGCCAGAAAGCTGCGAGAGGTAACTCTCCAGACAGGAATTAAGAGG 483  
Db  
361 TCTGCAGAGTGGCCAGAAAGCTGCGAGAGGTAACTCTCCAGACAGGAATTAAGAGG 420  
Qy 484 TTTGTGCGACGGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCTCCATCTGTGGAA 543  
Db  
421 TTTGTGCGACGGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCTCCATCTGTGGAA 480  
Qy 544 AGCCCGATTTATAGGAACATCAATTTGATGAGCTTGTCAAGCATACCAAGACAGCGCC 603  
Db  
481 AGCCCGATTTATAGGAACATCAATTTGATGAGCTTGTCAAGCATACCAAGACAGCGCC 540  
Qy 604 AAAGGACTTCTGGATGGCGGGTGTGATCTTACTCATGAACTATTTTGTACTGCCC 663  
Db  
541 AAAGGACTTCTGGATGGCGGGTGTGATCTTACTCATGAACTATTTTGTACTGCCC 600  
Qy 664 AATGCCAAGCGAGCTTTGTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG 723  
Db  
601 AATGCCAAGCGAGCTTTGTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG 660  
Qy 724 CCTATCTTTATTCAGGACGATGTTGTGATAAAAGTGGCGGACTCTTTCGGACAGACA 783  
Db  
661 CCTATCTTTATTCAGGACGATGTTGTGATAAAAGTGGCGGACTCTTTCGGACAGACA 720  
Qy 784 GGAGAGGATTTGTGTCATCAGGCTCTCATGGAGAACCACTCTGTCATGGAATTAATGTT 843  
Db  
721 GGAGAGGATTTGTGTCATCAGGCTCTCATGGAGAACCACTCTGTCATGGAATTAATGTT 780  
Qy 844 GCTTTGGGTGCGAGCTGAGATGAGACCTTTTATTGAAAATAATTTGAAAATGTACACAGCC 903  
Db  
781 GCTTTGGGTGCGAGCTGAGATGAGACCTTTTATTGAAAATAATTTGAAAATGTACACAGCC 840  
Qy 904 TATGTCCTCTGTTATCCCAATGCAAGTCTTTCACCAACACTTTTGGTGACTATGATGAACG 963  
Db  
841 TATGTCCTCTGTTATCCCAATGCAAGTCTTTCACCAACACTTTTGGTGACTATGATGAACG 900  
Qy 964 CTTCTATGATGGCCAAAGCACCTTAAGGATTTTGTCTATGATGCTGGTCAATATAGTT 1023  
Db  
901 CTTCTATGATGGCCAAAGCACCTTAAGGATTTTGTCTATGATGCTGGTCAATATAGTT 960  
Qy 1024 GGAGGATGCTGTGGTCAACACAGATCATATCAGGAAATTTGTAAGCTGTGAAAAAT 1083  
Db  
961 GGAGGATGCTGTGGTCAACACAGATCATATCAGGAAATTTGTAAGCTGTGAAAAAT 1020  
Qy 1084 TGTAAGCCTAGAGTTTCCACTGCTGCTTTTGAAGGACATATGTTTACTGCTGGTCTA 1143  
Db  
1021 TGTAAGCCTAGAGTTTCCACTGCTGCTTTTGAAGGACATATGTTTACTGCTGGTCTA 1080  
Qy 1144 GAGCCCTTCAGGATGAGCCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAATGTT 1203  
Db  
1081 GAGCCCTTCAGGATGAGCCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAATGTT 1140  
Qy 1204 GCAGGATCAAGGAAGTTTGTCTAAACTCATCATGGCAGGAAACTATGAAGAGCCTTGTGT 1263

1141 GCAGGATCAAGAGTTTCTTAATCTATCATGCGAGGAATCTATGAAGAAGCTTGTGT 1200  
1264 GTTCCAAAGTCAGGTGGAAATGGAGCCAGGTGTTGGATGTCACATGATGATGGC 1323  
1201 GTTCCAAAGTCAGGTGGAAATGGAGCCAGGTGTTGGATGTCACATGATGATGGC 1260  
1324 ATGCTAGATGTCAGAGTCAGATGCAATGTTTGCACATTAATGTTCCGAGCCAGAC 1383  
1361 ATGCTAGATGTCAGAGTCAGATGCAATGTTTGCACATTAATGTTCCGAGCCAGAC 1320  
1384 ATCCAAAGGTACCTTGTGTCATGCACTCCCTCCAAATTTGCTGATGAAGCTGGGTTA 1443  
1321 ATCCAAAGGTACCTTGTGTCATGCACTCCCTCCAAATTTGCTGATGAAGCTGGGTTA 1380  
1444 AAGTCTGCCAAGGAAGTGCATTTGTCATAGCAATTAAGTCTGAAGGAAGGAGGACGAC 1503  
1381 AAGTCTGCCAAGGAAGTGCATTTGTCATAGCAATTAAGTCTGAAGGAAGGAGGACGAC 1440  
1504 TTCTTGGAAGAGCCAGGAGATTAAGAAAGTATGAGCTGCTATGTTGCTCATGGCTTTT 1563  
1441 TTCTTGGAAGAGCCAGGAGATTAAGAAAGTATGAGCTGCTATGTTGCTCATGGCTTTT 1500  
1564 GATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCTAC 1623  
1501 GATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCTAC 1560  
1624 CATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATATC 1683  
1561 CATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATATC 1620  
1684 CTAACCATTTGGGACTGGATGAGGAAACACAACTTGTATGTCATTAATTTATCCATGCA 1743  
1621 CTAACCATTTGGGACTGGATGAGGAAACACAACTTGTATGTCATTAATTTATCCATGCA 1680  
1744 ACAGAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAGCTTTTCCAACTTG 1803  
1681 ACAGAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAGCTTTTCCAACTTG 1740  
1804 TCCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTTCCTTTAC 1863  
1741 TCCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTTCCTTTAC 1800  
1864 CATGCAATCAAGTCTGGCATGGAACATGAGATAGTGAATGCTGGAACCTCCCTGTGTAT 1923  
1801 CATGCAATCAAGTCTGGCATGGAACATGAGATAGTGAATGCTGGAACCTCCCTGTGTAT 1860  
1924 GATGATATCCATAGGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCTT 1983  
1861 GATGATATCCATAGGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACCTT 1920  
1984 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGACAGAGGGAAGAAAGTC 2043  
1921 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGACAGAGGGAAGAAAGTC 1980  
2044 ATTGAGCTGATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAATGCTGCTTGTG 2103  
1981 ATTGAGCTGATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAATGCTGCTTGTG 2040  
2104 AAGGCCATTGAAACATATTTATGAGATGATGAGGAGCCAGGTTAAACCAAAAAA 2163  
2041 AAGGCCATTGAAACATATTTATGAGATGATGAGGAGCCAGGTTAAACCAAAAAA 2100  
2164 TATCCCGACCTCTCAATATTAATTAAGAGCCCTGATGAATGAATGAATTTGTTGGT 2223  
2101 TATCCCGACCTCTCAATATTAATTAAGAGCCCTGATGAATGAATGAATTTGTTGGT 2160  
2224 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGGTTATG 2283  
2161 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGGTTATG 2220  
2284 AAGAGGCTGTTGGCCACTTATCCCTTTCATGGAATAAGAAAGAGAAACAGAGTG 2343

2221 AAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAATAAGAAAGAGAAACAGAGTG 2280  
2344 CTTTAAAGGCACTAGAGAAAGAGGACCTTACAGAGGACCAATCGTGTGGCCACTGTTT 2403  
2281 CTTTAAAGGCACTAGAGAAAGAGGACCTTACAGAGGACCAATCGTGTGGCCACTGTTT 2340  
2404 AAAGGGCAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGTGGCTCAATAAT 2463  
2341 AAAGGGCAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGTGGCTCAATAAT 2400  
2464 TTCCGAGTTTATGATTTAGGAGTCATGACTCCATCTGTAAGGATACCTGAAAGCTGCTCTT 2523  
2401 TTCCGAGTTTATGATTTAGGAGTCATGACTCCATCTGTAAGGATACCTGAAAGCTGCTCTT 2460  
2524 GACCAAAAGCAGATTAATTTAGGCTGTGAGGACTCATCTCTCCCTCCCTGGAATGAATG 2583  
2461 GACCAAAAGCAGATTAATTTAGGCTGTGAGGACTCATCTCTCTCCCTGGAATGAATG 2520  
2584 ATTTTGTGTCGCAAGGAATGGAGAGATAGCTATAGGATTCATTTGTTGATTTGGAGGA 2643  
2521 ATTTTGTGTCGCAAGGAATGGAGAGATAGCTATAGGATTCATTTGTTGATTTGGAGGA 2580  
2644 GCAACCACTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTGCACCTGTA 2703  
2581 GCAACCACTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTGCACCTGTA 2640  
2704 ATCCATGCTCTGACGCGTCCAAAGAGTGGTGGTGTGTTCCAGCTGTTAGATGAATAAT 2763  
2641 NNCCATGCTCTGAGCAGCTCCAAAGAGTGGTGGTGTGTTCCAGCTGTTAGATGAATAAT 2700  
2764 CTAAGAGTGAATTAATTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCAT 2823  
2701 CTAAGAGTGAATTAATTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCAT 2760  
2824 TATGAGTCTCTCAAGAGAGGAGATTAATTAAGCTTAAAGTCAAGCCAGAAAAAGTGGTTTC 2883  
2761 TATGAGTCTCTCAAGAGAGGAGATTAATTAAGCTTAAAGTCAAGCCAGAAAAAGTGGTTTC 2820  
2884 CAAATGATGTTGGCTGTCTGAACTCTCAAGAGTGAAGCCAGCTTATTTGGGACCCAGGTC 2943  
2821 CAAATGATGTTGGCTGTCTGAACTCTCAAGAGTGAAGCCAGCTTATTTGGGACCCAGGTC 2880  
2944 TTTGAGAGCTATGACCTGCAAGAGCTGGTGGACTACATTTGACTTGGAGCCCTTCTTTGAT 3003  
2881 TTTGAGAGCTATGACCTGCAAGAGCTGGTGGACTACATTTGACTTGGAGCCCTTCTTTGAT 2940  
3004 GTCTGGCAGCTCCGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAACGACAAA 3063  
2941 GTCTGGCAGCTCCGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAACGACAAA 3000  
3064 ACAGTAGTGGAGAGGCGCAGGAAAGGTCTACGATGATGCCCAATATGCTGAAACACACTG 3123  
3001 ACAGTAGTGGAGAGGCGCAGGAAAGGTCTACGATGATGCCCAATATGCTGAAACACACTG 3060  
3124 ATTAGTCAAAAGAAACTCCGGGCGCGGGTGTGGTGGTGTGGCCAGCAGCAGAGTATC 3183  
3061 ATTAGTCAAAAGAAACTCCGGGCGCGGGTGTGGTGGTGTGGCCAGCAGCAGAGTATC 3120  
3184 CAAGACGACATTCACCTGTACGAGAGGCTGTGTCGCCAGGCTGCAGAGCCCATAGCC 3243  
3121 CAAGACGACATTCACCTGTACGAGAGGCTGTGTCGCCAGGCTGCAGAGCCCATAGCC 3180  
3244 ACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGGAGCCATATAC 3303  
3181 ACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGGAGCCATATAC 3240  
3304 TGCTCTCAGACTTCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3363  
3241 TGCTCTCAGACTTCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3364 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3423  
3301 GCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360



Db 2507 ACATGGCAAGGCAATGCTGACCTGCAAGCGGCCATGGTCACTTCTCAACCTGATCG 2448  
Qy 1370 CTTCCGAGCAGACATCGCAAGGTAACCTTTGTGATGACGACTCTCTCAATTTTCTGTGA 1429  
Db 2447 CTTCCGAGCAGACATCTCGCGCTGCGGATCATGATCATGATCATGATCATGATCATGAT 2388  
Qy 1430 TTGAAGCTGGGTAAAGTGTGCGCAAGGGAAGTCAATTTGCAATGATGATGATGATGATG 1489  
Db 2387 TCGAGCGCGGCTGAGTGTGATTCAGGCAAGGGAATGCTCACTGATCTCGATGATGATG 2328  
Qy 1490 AAGGAGGAGCAGCTTTCTGGAGAAGGCCAGGAAGATTTAAAGATATGAGGCTCTATGG 1549  
Db 2327 AAGCGCTCGAGGCTTCAAGCACCATGCCGCTGCAAGCGCTACGGCGCGCGGTGG 2268  
Qy 1550 TGGTCATGGCTTTGATGAGAAGGACAGGCAACAGAAACAGACACAAAATCAGAGTGT 1609  
Db 2267 TGGTGTGGCTTCGACGAGGACGCGCCAGGCCAGCACCCAGGCGCGCAAGGAATCT 2208  
Qy 1610 GCACCGGGCTTACCCTCTGCTGTGAAATACTGGGCTTTAATCCAAATGACATATTT 1669  
Db 2207 GCAAGCGCTCTACGACATCTGCTGACGAGTGGCTTCCCAACCGAAGACATCATCT 2148  
Qy 1670 TTGACCTTAATATCTTAACCATTTGGACTGGAAATGGAGAACACAACTTGTATGCCATTA 1729  
Db 2147 TCGATGCGAATCTTCCGCAATCGCCACCGCATCGAGAACACAACTACGCGGTG 2088  
Qy 1730 ATTTTATCCATGCAACAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAG 1789  
Db 2087 ATTTTATCCATGCAACAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAGTGGAG 2028  
Qy 1790 GTCTTTCCAACTTGTCTTCTCTTCCGAGAAATGGAGCAATTCGAGAGCAATGATG 1849  
Db 2027 GGGTGTCCAACTGCTCTTCTCTTCCGAGCAACAAACCGGTGCGGAGCGATCCACT 1968  
Qy 1850 GGGTGTCTTACCATCAATCAAGTCTGCGATGAGCATAGTGAATAGTCTGAA 1909  
Db 1967 CGGTGTCTTACTACTACGATCGCAACGCGCTGACATGGGATGCTCAACGCGGCC 1908  
Qy 1910 ACCTCCCTGTGATGATATCAATAGGAATCTTCTGACGCTCTGTAAGATCTCATCT 1969  
Db 1907 AGCTGGAATCTACGACGAGATTCGGAAGCGCTGCGCACCGGCTCGAGACGTGTGC 1848  
Qy 1970 GGAATAAGACCTTGAGGCACTGAGAGCTCTTAAGTTATGCCAGACTCAA----GGCA 2026  
Db 1847 TCAAACCGCACCGCGAGGCCACCGAGGCCCTGCTGGCGATGCGCGAGCACTACAAGGGCG 1788  
Qy 2027 CAGGAGGGAAGAAAGTCAATCAGACTGATGATGAGTGAAGAAATGGCCCTGCGAAGCGCC 2086  
Db 1787 CGCGCGGCTCAAGAGGCGGAGGACGAGGAATGGCGAGCTACAGTGTGAGAGCGCC 1728  
Qy 2087 TTGAGTATGCCCTTTGTAAGGCGCATTTGAAACATATATTATGAGGATCTGAGGAAGCCA 2146  
Db 1727 TCGAGCATGCGTGTCAAGGCGATCACCACTGGATCGTGAAGCAACCGAGA----- 1673  
Qy 2147 GGTAAACCAAAAAATATCCCGAAGCTCTCAATATAATTTGAAGGACCCCTGATGAATG 2206  
Db 1672 ----ATGCGCGCAGCAGTGTGCGGCTCCATCGAGGTCATCGAAGGTCCGCTGATGTCCG 1617  
Qy 2207 GAATGAAATTTGTTGTGATCTTTTTCGAGCTGGAATAATGTTTCTACCTCAGGTTATAA 2266  
Db 1616 GGATGAACGTGTGCGGACCTGTTTCGCGCGCGCAAGATGTTCTCCCGCAGGTGTCA 1557  
Qy 2267 AGTCAGCGCGGTTATGAAGAGGCTGTGGCCACCTTATCTCTTTCATGGAAGAAAGAA 2326  
Db 1556 AGTCGCGGAGTGTGATAACAGGCGGTGGCCACCTGATTTCTCTTCTATCGAGCGGAGA 1497  
Qy 2327 GAGAAGAAACCAAGTGTCTTAACGGGCAAGTGAAGGAAGGACCCCTTACCGAGGCCACA 2386  
Db 1496 AAGCGGATA-----AGCCGGAAGCCAAAGGGCAAGA 1467  
Qy 2387 TCGTGTGGCCCATGTTTAAAGCGCAGCTGACGACATAGGCAAGAAATAGTGTGAGTAG 2446  
Db 1466 TCCTGATGGCCACCGGTGAAGGGCGACGTGACGACATCGGCAAGAAATCGTGGCGGTGG 1407

Qy 2447 TCCTTGGCTGCAATAATTTCCGAGTTATTGATTTAGGAGTCAATGACTTCCATGTGATAGA 2506  
Db 1406 TGCTCGGCTGCAACCGGCTATGACGTGTGTCACCTCGGCGTGTGTCGCGGAGAGA 1347  
Qy 2507 TACTGAAAGCTGCTCTTGACCAACAGCAGATATAATTTGGCCCTGTCAAGACTCATCACTC 2566  
Db 1346 TCTGACAGCCGATCGCGAGAAATGCGACATCATCGGCCCTGTCTGGCTGATCAAGC 1287  
Qy 2567 CTTCCCTGGATGAATGATTTTGTTCGCAAGAAATGAGAGATTAAGCTATAAGGATTC 2626  
Db 1286 CGTCTGGAGAGATGTTCCACGTGCGCAAGAAATGACGCGCAGAAATTCAGTTGC 1227  
Qy 2627 CATTTGTTGATGGAGGAGCAACCACTTCAAAAACCCACAGCAGTTAAATAGCTCCGA 2686  
Db 1226 CGTGTGATGATCGGCGCGCCACTACCTCGAAGCGCATACCGCGGTGAAGATCGATCCGC 1167  
Qy 2687 GATACAG-----TGACACCTGTAAATCATGTCTCTGACGCGTCCAAAGAGTGTGTGTGT 2743  
Db 1166 AGTACAGCAACGACGCGGTGTTCTACGTACCGACGCTCGCGCGGTAGCGTGCCCA 1107  
Qy 2744 CCAGCTGTTAGATGAATAATCTAAAGGATGAATACTTTGAGGAAATCATGGAAGAAATG 2803  
Db 1106 CCAGCTGCTGCTCAAGGAGCTGAAGCGGACTAGCTGGCCCGCACCGCGCGGACTACG 1047  
Qy 2804 AGATAT-----TAGACAGGACCAATTATGATCTCTCAAGAGAGGAGATCTTACCCCTT 2858  
Db 1046 CGGTGTGCGGCAACGCAACCGCGCAACCGCAGCGCCGACCGAGCGGCTGAGCTACCAAC 987  
Qy 2859 AAGTCAAGCAGAAAAGTGGTTTCCAAATGGAATGGCTCTCTGAACTCACCCAGTGAA 2918  
Db 986 AGCGGATCGCAACAGCGCGGCTTGGACTGGCGCGGCTACCAAGGCGCGGAC----- 935  
Qy 2919 GCCACGTTTATTTGGGACCCAGGCTTTTGAAGACTATGACTTCAGAAAGCTGTGTGACTA 2978  
Db 934 GCCTTCTTACCGCGGTGAGGCTGCTCGAGAGATCGACTCGCGGTGTTCGCCGAGTA 875  
Qy 2979 CATGACTGGAAGCCTTTCTTGTATGCTGTGGAGCTCCGGGCAAGTACCGAATCGAGG 3038  
Db 874 CATGACTGGAAGCCTTTCTTCTTCTTCTGGAACCTTGGCGGCAAGTACCG----- 823  
Qy 3039 CTTCCCCAGATATTTAAAGCAAAACAGTAGTGGAGGCGCAGGAAGGCTCTACGATGA 3098  
Db 822 -----CGCATCTCACCGAAGTGTGCGGAGGCGCCCACTCTGTTGTTCAACGA 770  
Qy 3099 TGCCCAATATGCTGAACACTGATTTAGTCAAAAGAAACTCGGCGCCGGGTGTGGT 3158  
Db 769 CGCCAGCGGCTGTGAGAGCTGATCGAGAGAGCTGATCAAGGCCCCGCGGTGTT 710  
Qy 3159 TGGGTTCTGGCCAGCAGAGTATCCAAGACGATTTCACTGTACGAGAGGCTGCTGT 3218  
Db 709 CGGCTTCTGGCGGCCAACCGAGTCCGACACGACCTGGAGGTCTACGCGCGCGATGG 650  
Qy 3219 GCCCAGGCTGAGAGGCCATACGACTTCTATGGGTTAAGGCAACAGGCTGAGAGGA 3278  
Db 649 C-----GAGACCTCGCCACCTTGCACCTCGCGCAGCAGATCAAGCC 602  
Qy 3279 CTCTGCGCAGCAGGAGCCTACTGCTCTCAGACTTCTCATCTGCTCTTGTGATTTCTGG 3338  
Db 601 GGAAGGCA-----AGCCGAACTGTGCTGGCGGATTTCTGCGCGCGAAGAAAGCGG 548  
Qy 3339 CATCCGCTGACTACTCGGCGCTGTTTG-----CCGTTGCTGCTTTGGGGTAGAAGAGCTGAG 3395  
Db 547 CGTGGCGGACTCATCGGCGCTTCTATCAACCGCGCGGATCGGCGCGGAGGAGTGGC 488  
Qy 3396 CAAGGCTTATGAGGATGATGCTGACGATACAGCAGGATCATGCTGAGCGCTGAGGGA 3455  
Db 487 CAAGGCTTATGAGGATGATGCTGACGATACAGCAGGATCATGCTCAAGGCGCTCGCGCA 428  
Qy 3456 CCGCTGCGCAGGCGCTTTGCAAGAGGCTTCCATGAAGAGTTCGCGCGAAGCTGTGGGC 3515  
Db 427 CCGCTGCGCGCAAGCTTGGCGCGGCTGCAAGAGGCTGCGGCGGCTGCGCAAGGAGTACTGGGG 368



QY 3516 CTACTGTGGCAGTGAAGCAGTGGAGCTCGCAGACCTCGGAAGTTGGCGTCAAGGGCAT 3575  
Db |||||  
QY 367 CTAGCCCGCAGCAGCAGCTCGACAGAGCCCTTGATCAAGGCAATACGTCCGCAT 308  
Db |||||  
QY 3576 CGGCCGGCTCTGTGCTACCCAGCAGCCGCGACACACCGAGAGCTCACCATTGTGGAG 3635  
Db |||||  
QY 307 CGGCCGGCAGCAGCGCTACCCGGCTGCGCCGACCATACCGAGAAAGCACCCTGTTGGA 248  
Db |||||  
QY 3636 ACTCGCAGACATCGAGCAGTCTACAGGCATTTAGGTTAACAAATCATTTAGCAATGGCAC 3695  
Db |||||  
QY 247 ACTGCTGATCCGAGGCGCTGTGCGGCTGAGCCGTGACCGAGCACTACGCGATGTTCCC 188  
Db |||||  
QY 3696 TGCTTACAGCAGTCTCAGGCGCTTACTCTTCCAAATTTGAAGTCCAAATATTTGCTGTGG 3755  
Db |||||  
QY 187 CGCCGGCGCTGAGCGGTTGGTATTTGCGCCACCGCAGCGCAGTACTTTCGCGGTGG 128  
Db |||||  
QY 3756 GAAGATTTCCAGGATCAGGTTGAGGATTTATGCAATGAGGAAGAACATATCTGTGGCTGA 3815  
Db |||||  
QY 127 CAAGATCGACAAGGACCAAGGTGGAACGCTACAGCCAGCGCAAGGGCCAGGAAGCCAGCGT 68  
Db |||||  
QY 3816 GGTGAGAAATGGCTTGGACCCATTTTGGGATATGAT 3852  
Db |||||  
QY 67 CAGGAGCGCTGGCTGGCGCGCAACCTTGGCTACGAT 31  
Db |||||

## RESULT 11

US-09-252-991A-13693  
; Sequence 13693, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 13693  
; LENGTH: 3795  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13693

Query Match 23.0%; Score 903; DB 3; Length 3795;  
Best Local Similarity 55.2%; Pred. No. 1.9e-265;  
Matches 2074; Conservative 0; Mismatches 1580; Indels 103; Gaps 12;

QY 113 CCCTCGGGATGAGATCAATGCCATTCTGCAAGAGGATTTATGCTGTGATGGAGGA 172  
Db |||||  
QY 119 CCGGCTGCAAGCCCTCCAGCAGCGCCCTCAGGGAACGATATCTGATCTCGATGCGGCA 178  
Db |||||  
QY 173 TGGGACCATGATCCAGCGGAGAGCTTAAACGAAGACACTTCCGAGCTCAGGAATTTA 232  
Db |||||  
QY 179 TGGGACCATGATCCAGAGCTCAAGCTGGAAGAGCGGACTTACC CGCGGCGGCGCTTCG 238  
Db |||||  
QY 233 AAGATCATGCCAGCGGCTGAAAGGCAACAAATTTTAAAGTATAACTCAGCCCTGATG 292  
Db |||||  
QY 239 CCGACTGCGCGAGCGATGGAAGGCAACAGCACTCTTGCTGTGAGCCCGCGAGC 298  
Db |||||  
QY 293 TCATTTACAAATCCATAGGAATTAATCTGCTGGCTGGGCGAGATATCAATTTGAAACAAATA 352  
Db |||||  
QY 299 TGATCCAGGCCATCGAAGAGGCTTACCTCGACCGCGGCGCGACATCTCTCGAGACCAACA 358  
Db |||||  
QY 353 CTTTATGAGCAGCTAGTATTCGCCAAGCTGACTATGGCTTGAACACTTGGCTTACCGGA 412  
Db |||||  
QY 359 CCTTCAACGCCACCAGGCTGCTCCAGGCGGACTACGGATGCGATGCGCTGGCTTACGAAC 418  
Db |||||  
QY 413 TGAACATGTCTCTCAGGAGTGGCCAGAAAGCTGCCGAGAGGTAACCTCTCCAGA--- 469  
Db |||||

Db 419 TCAACCTCGAAGGGGCGCGCTTGGCCGCGAGGTGGCGGAGCGAAGACCGCCGAGACCC 478  
QY 470 CAGGAATTAAGAGGTTTGTGGCAGGGGCTCTGGGTCGAGCTAATAAGACACTCTCTCTGT 529  
Db 479 CGGACAAAGCCGCGCTTGTGCGCGGTGTCTGGCCGAGCAGCCGCACTCTGCTCGATTT 538  
QY 530 CCCATCTGTGGAAGGCGCGGATTTATAGGAACATCACTTTGATGAGCTTTGTTGAAGCAT 589  
Db 539 CCGGAGCGTGAACAACCCCGCTACCGCAACCTGCTCTTCGACGAACTGCTGGAGAACT 598  
QY 590 ACCAAGAGAGGCGCAAGAGACTCTTGGATGCGGGGTTGATATCTTACTATTGAAACTA 649  
Db 599 ACGTGAGGCGGACCGCGGCTGATCGAAGGCGCGCGACCTGATCTGATCGAGACCA 658  
QY 650 TTTTGTGATATGCGCAATGCGAAGCAGCTTGTGTCCTCCAAATCTTTTGTAGGAGA 709  
Db 659 TCTTCGACACCTCAAGCCCAAGCGCGGATCTTGGCGGTCAGGGGCTGTTGAGGAAC 718  
QY 710 AATATGCTCCCGGCTTATCTTTATTCAGGAGCGATCGTTGTGATAAAGTGGGCGGACTC 769  
Db 719 TCGCGGTGGAGCTGCGGATCATGATCTCGGAACCATCACCGACGCTCCGCGCGCACCC 778  
QY 770 TTTCCGACAGACAGGAGAGGATTTGTCTATCAGCGGTGTCTCATGGAAGAACCTCTGCA 829  
Db 779 TGTGCGGCCAGACCAAGGCGCTTCTGGAACCTCGGTGCGGCATGCGCGCGGCTCTCGG 838  
QY 830 TTGGATTAATTTGCTTTGCGGTGCGAGCTCAGATGAGACCTTTTATTTGAAATTAATTCGAA 889  
Db 839 TAGGCTGGAATGCGGCTCGGCGCCAAAGAAATTGCGCGGTATCATCGAGGAACCTGTGCA 898  
QY 890 AATGTACAACAGCCTATGCTCTCTGTTATCCCAATGAGGCTTCTCCCAACACCTTTGGTG 949  
Db 899 CCAAGGCGGACACTCACTGCTCGGCCCAACCGCGCGCTTCCCGAAGCGCTTCCGCG 958  
QY 950 ACTATGATGAAACCGCTTCTATGATGCGCAAGCACTAAAGGATTTTGTCTATGATGGCT 1009  
Db 959 AATACGAGCAATCGCCGCGGAAATGCGCGGTGTGTCGAGGAATTCGCGCGCGCGCT 1018  
QY 1010 TGGTCAATATAGTTGGAGGATGCTGGGTCAACACAGATCATATCAGGGAATTCCTG 1069  
Db 1019 TCCTCAATATGCTCGCGGCTGCTGCGGCAACCGCGGCGCATCGAGCGGATCGCCA 1078  
QY 1070 AAGCTGTGAAATTTGAGCTAGAGTTCCA CCGTCCACTGCTTTTGAAGGACATATGT 1129  
Db 1079 AGGAGTGGCAAGTACC CGCGCGGCCCATCCCGGAGATTTCCCGCGGCTGTGTC----- 1132  
QY 1130 TACTGTCTGTCTAGAGCCCTTCAAGATTGGAACGTTACCAACTTTTGTAAAGATTGGAG 1189  
Db 1133 GCCTGTCCGGCTGGAGCGGTTCAACCATCGACCGAGCTCGCTGTTCGTCAACGTCGCGG 1192  
QY 1190 AGCGCTGTAATGTTGCAAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGAAACTATG 1249  
Db 1193 AGCGCAACCAATCACCGGTTCCGCGCAAGTTCCGCGCGTGTATCCGGAGGAATCTACG 1252  
QY 1250 AAGAGCGCTTGTGTTGTTGCCAAAGTGCAGAGTGGAAATGGGAGCCAGGTTTGGATGTCA 1309  
Db 1253 CGGAAGCTCTCGAGGTCGCCAGCAGCAGGTGGGAAGCGCGCGCCAGGTATCGACATCA 1312  
QY 1310 ACATGATGATGCGATGCTAGATGGTCCAAAGTCAATGACAGATTTTGGCAACTTAATTG 1369  
Db 1313 ACATGAGCAAGGATGCTGAGCTCGAAGGCGGCGCATGTCACCTTCTCAACCTGATCG 1372  
QY 1370 CTTCCGAGCCAGACATCGCAAGGTAATCTTTGTGATCGATCGCTCCAAATTTTCTGTGA 1429  
Db 1373 CCTCCGAGCCGACATCTCGCGCTGCGGATCATGATCGACTCTCTCAAGTGGGAAGTGA 1432  
QY 1430 TTGAAGCTGGGTTAAAGTGTCCCAAGGAAGTGCATTTGTCAATAGCATTAAGTCTGAAGG 1489  
Db 1433 TCGAGGCGGCTGAGTGATCCAGGCAAGGCGATCGTCAACTCGATCTCGATGAAGG 1492  
QY 1490 AAGGAGGAGCAGCTTCTTGGAGAAAGCGAGGAAGATTTAAAGATATGAGAGCTGTATGG 1549  
Db 1493 AAGGCTCGAGGCTTCAAGCAACATGCGCGCTGTGCAAGCGCTACGCGCGCGGCTGG 1552





Db	3633	CGCCGCGCGGTTCAGCGTTCGTTGGTATTTGCGCCACCCGAGCGCAGTACTTCGCGGTGG	3692
Qy	3756	GAAGATTTCAAGGATCAGGTTGAGGATTATGCATTGAGGAAGAAACATACTGTGGTGA	3815
Db	3593	CNAGATCGACAAAGACCAAGTGGAACTACAGCCAGCGCAAGGGCCAGGAAGCCACGT	3752
Qy	3816	GGTTGGAATATGGCTTCGACCCATTTGGGATATGAT	3852
Db	3753	CAGCGAGCGTGGCTGGCGCGCAACCTTGGTTCAGT	3789

## RESULT 12

```

US-09-543-681A-3077
; Sequence 3077, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3077
; LENGTH: 3750
; TYPE: DNA
; ORGANISM: proteus mirabilis
US-09-543-681A-3077

```

Query Match 23.0%; Score 901; DB 3; Length 3750;  
Best Local Similarity 55.0%; Pred. No. 7.9e-265;  
Matches 2045; Conservative 0; Mismatches 1575; Indels 99; Gaps 10;

Qy	145	AAGAGGATATTGGTGTCTGGATGGAGGATGGGGACCATGATCCAGCGGGAGAGCTAAAC	204
Db	112	AAACGTATCTTAGTCTCTTGTATGGCGCAATGGAAACCATGATCCAGCAATATCAGTTGGCA	171
Qy	205	GAAGAACACITCTCCAGGTCAGGAATTTAAAGATCATGCCAGGCCCTGAAAGGCAACAAT	264
Db	172	GAAGAAGATTACCAGGTGAACGTTTGGCCCAITGGCAATGTGATGTAAAGGGGAATAAT	231
Qy	265	GACATTTTAAGTATAACTCAGCCTGATGTCTATTTACCAATCCATAAGGAATACTTGGCTG	324
Db	232	GACCTTTTAGTCTTGAACGAGCCGAGATCAATTACTGNAATTCATATGCTTATTTTGA	291
Qy	325	GCTGGGCGAGATATCATTTGAACAAATACTTTTAGCAGCACTAGTATTGCCCAAGCTGAC	384
Db	292	GCCGGTGCAGATATCGTAGAACAACAAACACCTTTAATGCAACCTCTATTGGCCATGGCTGAT	351
Qy	385	TATGGCCTTTGAACACTTGGCCCTACCGGATGAAACATGTGCTCTGCAGAGATGGCCAGAAAA	444
Db	352	TATCAGATGGAAAGGGTTATGTGCAGAAATTTGAATGAAGAGGCTGTAAACTTGGCCAGAGCC	411
Qy	445	GCTCCCGAGGA--GGTAACTCTCCAGACAGGAATTAAAGAGTTTGTGGCAGGGGCTCTG	501
Db	412	TGTGCTGATAAATGGAGTGCATTAACGCCAGATAAACACAGTTATGTGTGACGGGGTATTA	471
Qy	502	GGTCCGACTAATAAGACACTCTCTGTGTCCCCTCTGTGGAAAGCCGGGATTTATAGGAAC	561
Db	472	GGGCCAACCAATAGAACAGCATCCATTCTCTGATGTTAACGATCCCGCATTTTCGTAAT	531
Qy	562	ATCATTGATGATGCTTGTGTAAGCATACCAAGACAGGCCAAAGACCTTCTGGATGGC	621
Db	532	ATCTCATTTGATACATTAGTATGGCTTATCGAGAGGCGATCTCGGCACTCATTTAAAGGA	591
Qy	622	GGGGTGTGATCTTTACTCATTTGAACATAATTTTTCATCTGCCAATGCCAAGGCGACCTTG	681
Db	592	GGGGTGGATTTAAATATGGTGGAAACCAATTTTGTACCTCTAAATGCTAAAGCGCGCTATT	651
Qy	682	TTTGCACTCCAAAATCTTTTTTGAGAGGAAATATGCTCCCGGCGCTATCTTTTATTTTCAGGG	741

1726	CAACTGCGCTTAATGCTTTAAATTTCTGGCGGGTGTCTTAATGTGTCATTTCTCTTTTCGGGGC	1781
1822	ATGGAAGCCATTTCGAGAAGCAATGCATGGGGTTTTCTTTTACCATGCAATCAAGCTCTGGC	1881
1786	AATGATCCCGTTTCGGAAGCGATCCACTCCGTTTTTCTCTATTACGCTGTAAAAAATGGT	1845
1882	ATGACATATGGAGATAGTGAATGCTTGGAAACCTCCCTGTGTATGATGATATCCATAGAAGAA	1941
1846	ATGGATATGGGATTTGTCAATGTCTGGCACTCGCTATTTATGACTCATTTAGCGGATGAA	1905
1942	CTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCTCCTGAGGCCACTGTAGAAGCTC	2001
1906	CTACGTAATCGGGTAGAGATGTTATTTTAAATCGCCATGCAGAGAGTACTGATAATTTTA	1965
2002	TTACGTTTATGCCAGACTCAAGGC-----ACAGGAGGGAAGAAAGTCATTTACAGACTGAT	2055
1966	CTGSCACTCGCAGACGCTATCTGTTGTACTAAAGTCAGGAAACAGCACTCTACGCTTGGC	2025
2056	GAGTGGAGAAATGCGCCCTGTGGAAGAACGGCTTGAGTATGCCCTTGTGGAAGGGCATTTGAA	2115
2026	CAATGGCGCAAAATGGGAGGTGGAAGAGCGCTAGAGTACGCGTTAGTGAAGGGGATCACC	2085
2116	AAACATATATTGAGGATACTGAGGAGCCAGGTTAAACCAAAAAAATATCCCCACCT	2175
2086	GAA'TTTTATCATTTGAAGATACAGAAG-----CGTGCGCCCAACAGCTTCGAGTCCT	2136
2176	CTCAATATAATTGAAGCACCCTCATGAATGGAATGAAAAATGTTGTGTGATCTTTTGGGA	2235
2137	ATCGAAGTGATGAAGGGCCATTAAATGAATGGCATGAACACAGTGGCGGATCTGTTTGGT	2196
2236	GCTGGAAAAATGTTTCTACCTCAGGTTTATAAAGTCAGCCCGGGTTATGAAGAAGCGTGTT	2295
2197	GAAGGTAAATGTTTTTACCCTCAGTAGTAAATCGGCACGGGTGATGAACAGGCCGCTA	2256
2296	GGCCACCTTATCCCTTTCATGGAAAAAGAGAAGAACCAAGAGTGTCTTAACGGCACA	2355
2257	GCTTACCTTGAACCTTATATTCAAGCACAATAAAAAAGCAGCAG-----	2300
2356	GTAGAAGAAGGAGCCCTTACCAGGCACCATCGTGTGGCCACTGTTAAAGGCGACGTG	2415
2301	-----CTGGCAGGAAAAAGTGTTACTTGGCGACAGTAAAGGGAGATGTC	2343
2416	CAGCATATGGCAAGAACATAGTTGGAGTAGTCCCTGGCTGCAATAAATTTCCGAGTTATT	2475
2344	CACGACATTGSCAAAAATATTGTTCGGAGTGTGTGCAATGTAATACTATGAAATTTATC	2403
2476	GATTTAGAGTCATGATCCCATGTGATAGAATACTGAAAGCTGCTCTTGACCAACAAGCA	2535
2404	GATCTGGGGTCAATGATGACCTTGCAGATAAAATCTTGCAAAACAGCCATTTGCAAAAATGTT	2463
2536	GATATAATTGSCCTGTCCAGGACTCATCACTCCTTCCCTGGATGAAATGATTTTGTGGCC	2595
2464	GATATAATTGGATTGTCTGGATTGATCACTCCATCACTGGAATGAGATGGTGAATTTGTGCT	2523
2596	AAGGAAATGGAGAGATTAGCTATTAAGGATTTCCATTGTTGATTTGAGGAGCAACCACTTCA	2655
2524	AAGGAGATGGAAGACCGGGTTTTCTCTTCTTAAATGATTTGGTGGTCGACAAATCG	2583
2656	AAAACCCAACACAGCAGTTTAAATAGCTCCGAGATACAGTGCACTGTGTAATCCATGTCTG	2715
2584	AAAGCGCACCGGCAGTAAAAAATTTGAACCTAAATTTACAGCCACCACCGTTTTATGTGCAG	2643
2716	GACCGGTCCAAGAGTGTGGTGTGTTCCACGTGTTAGATGAAATCTTAAAGGATGAA	2775
2644	AATCGGTCAAGAACCGTTCGGGGTGGTCCGTGCAATTTATCGGCCACACAAAAAGGCTGAT	2703
2776	TACTTTTGGAGAAATCATGGAAGAAATATGAAGATATTAGACAGGACCACTATGAGTCTCTC	2835
2704	TTTTGCTCGCTAAAAACACCGCGTGATATGAAGTGTGTAAGTCAAGCTATCGAGNAAAAAA	2763
2836	AAGGAGAGGAGATATTTACCTTTAAGTCAAGCCGAAAAAAGTGGTTTTCCAAATGGATTTGG	2895
2764	CCGCGTACGCTCCCGTTTTCTTTCAGCACAGCTCGTGCCTCAATGCATTCAGCTTGATCTGG	2823

QY	2896	CTGTCTGAACTCACCACAGTGAAGCCACCGTTTATTATGGGACCCAGGTCCTTTGAAGACTAT	2955
DB	2824	CAACACTATATAGCCACCTAAAGCCGAAACCAATATTAGGGGTACAACAAGTTACAGGCCAATATT	2883
QY	2956	GACCTGCAGAAAGCTGGTGGACTACATTGACTGGAAGCCTTTCTTTTGATGTCGCGAGCTC	3015
DB	2884	GAA-----ACTTTACGTGAGTATATCGACTGAGCGCTTTCTTTATGACTTGGTCATTG	2937
QY	3016	CGGGCCAAAGTACCCGAATCGAGGCTTCCCACAAGATATTTAACGACAAAAACAGTAGGTGGA	3075
DB	2938	CGCGGGAAATATCCACGT-----ATATTAGAAGATGAGGTGTGCGGAA	2982
QY	3076	GAGGCCAGGAAGGTCTACGATGATGCCCAACATATGCTGAACACACTGATTAGTCAAAG	3135
DB	2983	GAGGCAAGACGCGTTTTTGCAGATGCGAATGCCAATGTTGGATAAATCAAGCCGTGAAAA	3042
QY	3136	AAACTCCGGGCCGGGTGTGGTGTCTGGCCAGACACAGAGTATCCAAAGACGACATT	3195
DB	3043	TTACTGACACGAAAGGTATCGTTGGTTTATTCCTGCTAATCGCCTTGGTGATGACATT	3102
QY	3196	CACCTGTACCGAAGGCTGTGTGCCAGGCTGCAGAGCCCATPAGCCACTTTCTATGGG	3255
DB	3103	ATTATTTTACCAAGATGAACCGACACATG-----AGTTATTACGTGTGTGCCAT	3153
QY	3256	TTAAGGCAACAGGCTGAGAAAGGACTCTGCAGACACGGAGCCATPACTCTGCTCTCAGAC	3315
DB	3154	TTACGCCAACAAAAAGAAAAA-----GAGTACCCCTAACTATTTCCTTAGCAGAC	3204
QY	3316	TTCACTCGCTCCCTTGCAATCTTGGCATCCGTGACTACCTGGGCCCTGTTGGCTTGCCCTGC	3375
DB	3205	TTTATTGCCCTGTGGATAGTGGTCTTGTGACTATTTTGGTGTCTTTGAGTTACCGCGC	3264
QY	3376	TTTGGGGTAGAAGA--GCTGAGCAAGGCCCTATGAGGATGATGGTGACGACTACAGCAGC	3432
DB	3265	GGTCTAGAAGAAGATGCACCTCGCTAATGGGTATGATAATGCTCATGATGATTACAATAA	3324
QY	3433	ATCATGTCTAAGGGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAAGAAGCTCCATGAA	3492
DB	3325	ATTATGGTAAAGCGCTATCCGATCGTCTTTGCGAAGCTTTTGCAGAAATATTGTCATCAA	3384
QY	3493	AGAGTTCGCGAGAACTGTGGGCTACTGTGGCAGTGCAGCAGCTGGACGTCGACAGCCTG	3552
DB	3385	CAAGTGAGAACCAAAATTTGGGGTTACAGTCTCTGACGAAAGCGCTCTCTAATGACGAGTTG	3444
QY	3553	CGAAGGTTGCGGTACAAGGGCATCCGCCCTCTGGCTACCCACGCCAGCCGACCCAC	3612
DB	3445	ATTAGAGNAANAATCAAGGAACAAGCCTGCACCGGATATCCAGCTTGCCACGAGCAC	3504
QY	3613	ACCGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGCATTAGGTGA	3672
DB	3505	ACAAAAAAGCAAAAATTTGGCAACTGTTAAATGTAGAAAATCGCATTTGGTATGAGAGCTT	3564
QY	3673	ACAGAAATCATTAGCAATGGCACCTGCTTACGAGCTCTCAGGCCCTCTACTTCTCCNATTG	3732
DB	3565	ACTGATCCTTATGCGATGTGGCGGGGGCTTCGGTATCGGTTGGTATTTTATGTCATCCT	3624
QY	3733	AAATCCAAATATTTTGTGTGGGGAAGATTTCCAAGGATCAGGTTGAGGAATTATGCAATTG	3792
DB	3625	GAAAGTAATATTTTGGCGTCCGTCAAATTCAAAAGATCAGGTAGAGATTACGCCAA	3684
QY	3793	AGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTCGACCCATTTTGGGATATGA	3851
DB	3685	CGTCGTGGTATGAGTATGAGTGAAGTTGACGTTGGTTAGCCCTTAATTTAGGTTATGA	3743

RESULT 13

US-09-328-352-1578  
; Sequence 1578, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND

! TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

! FILE REFERENCE: GTC99-03PA

! CURRENT APPLICATION NUMBER: US/09/328,352

! NUMBER OF SEQ ID NOS: 8252

! SEQ ID NO 1578

! LENGTH: 3702

! TYPE: DNA

! ORGANISM: Acinetobacter baumannii

! US-09-328-352-1578

Query Match 22.9%; Score 895.8; DB 3; Length 3702;

Best Local Similarity 55.1%; Pred. No. 3.1e-263;

Matches 2060; Conservative 0; Mismatches 1567; Indels 114; Gaps 11;

Qy	128	TCAATGCCATCTCGAAGAGGATTAATGGTCTGGATGGAGGATGGGACCATGATCC	187
Db	35	TAAAAGCGCTCTTGGCTAAAGCATCTCTGATTAATGGTGAATGGGAACCATGATCC	94
Qy	188	AGCGGAGAACTAAACGAAGAACATCTCCAGGTCCAGGAATTTAAGATCATGCCAGGC	247
Db	95	AGCGCCATAAATTTGGAAGAGCTGACTATCGTGTGGAGCGTTTGGTGAATGGGCACATG	154
Qy	248	CGCTGAAAGGCAACAATGACATTTTAAAGTAAATCTCAGCCTGATGTCTATTTACCAATCC	307
Db	155	ACTTAAAGGTAAACATGACCTTTTGGTTTAAACAGCCTCAATCATTTCAAGGTATTC	214
Qy	308	ATAAGGAATACTTCTGCTGGCTGGGCGAGATATCATTTGAAACAAATCATTTTACGAGCTA	367
Db	215	ATGAAGCTACTCTCGATGCTGGTGCAGACATTAATGAAACCAACAGCTTTTAAACGGCAC	274
Qy	368	GTATGGCCCAAGCTGACTATGGCTTGAACACTTGGCTACCGGATGAACATGTGCTCTG	427
Db	275	GTGTTTCAATCTCTGACTATCACATGGAAGATCTGTTCCAGAGATTAAACCGTGAAGCAG	334
Qy	428	CAGAGTGGCCAGAAAGCTCCGAGGAGGTAACTCTCCAGACAGGAATTAAGAGGTTTG	487
Db	335	CACGTTTAGCCAAAGCAGCTTGGGAAA--ATATCGACTCCAGACAAACCGGTTTG	391
Qy	488	TGGCAGGGCTCTGGGTCGCACTAATAGACATCTCTGTGTCCCATCTGTGAAAGGC	547
Db	392	TGGCAGGTGTAATCTGGGGCCACATCTCGTATGTTCAATCTCTCCAGATGTAACCAACC	451
Qy	548	CGGATTATAGAAATCATATTCATGATGATGTTGTTGAGCATACCAAGCAGGCCAAG	607
Db	452	CTGCTTTTCGTAACATTAAGCTTTGATGAACATAAAGAAAATTTATATTGAAGCAACTCATG	511
Qy	608	GACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGATCTGCCAATG	667
Db	512	CACATAATTGAAGTGGTGCAGACATTAATCTGATTTGAACCTGATTTGATACTTTAACT	571
Qy	668	CAAAGCAGCCTGTGTTGCACTCCAAATCTTTTGTGAGGAGAAATATGCTCCCGGCCCTA	727
Db	572	GTAAGCAGCGATTTTTCAGTCAAGGAAGCTTTTAAACAAATTTGGTCCGGAATTAACCA	631
Qy	728	TCTTTATTTTCAGGACGATCTGTGATAAAGTGGCGGACTCTTTCGGACAGACAGAG	787
Db	632	TTATGATTTTCAGGGACCAATTAACGATGATCAAGGCGGACTTTAAACAGGTGAGACGGG	691
Qy	788	AGGATTTTGTATCAAGCGGTCTCATGAGAACCACTCTGCAATTTGGAATTAATTTGTGCTT	847
Db	692	AAGCTTTCTGGAATTCGTTCTCATGCGGATTTGCTTTCAATCGGTTTAACTGTGGCC	751
Qy	848	TGGGTGACGTGAGATGAGACTTTTATTTGAATAAATTTGGAATAATGTAACAAGCCTATG	907
Db	752	TTGGTGCAGATGCCATGCGCCTCACGTAAACAACTATTTCCGATGTCGACATACCTTTG	811
Qy	908	TCCTCTGTTATCCCAATCCAGGTCTTCCCAACACTTTTGGTGTGACTATGATGAAGCCTT	967
Db	812	TTTCAGCCCAACCAATTCAGGCTTTACCAACGCAATTTGGTGAATATGACGAATCCAG	871
Qy	968	CTATGATGGCCCAAGCACCTTAAAGGATTTTGTCTATGGAATGGTGGCTGCTCAATATAGTTGAG	1027

Db	872	AGCAAACTGCAGCTTTCTTAAAAAGAGTTGCTCGAAAGCGGTTTGAATTAACATTCACGGGTG	931
Qy	1028	GATGCTGTGGGTCAACACACAGATCATATCAGGGAATTTGCTGAAGCTGTCAAAAATTTGTA	1087
Db	932	GCTGCTGTGTGTACGACACACAGATCATATTCAGCTATTTGCAATTCGGGTAAAGACATTTG	991
Qy	1088	AGCTAGAGTTTCCACCTGCGCATCTTTTGAAGGACATATGTTACTGTCTGGTCTAGAGC	1147
Db	992	CGCTCGCCAAAGTGCTGAAACAGTACTGCTTGGC-----GCTTAAAGTGGTTTAGAAC	1045
Qy	1148	CCTTCAGGATTTGACCGGTACACCAACTTTGTTAACTTTGAGAGCGCTGTAAATTTGTGCAG	1207
Db	1046	CAITTAATAATTTATAATGATTCATTTGTTGTAACGTTGGTGAACGTAACGTTAACTACTG	1105
Qy	1208	GATCAAGGAATTTGCTAACTCATCATGCGCAGGAACATATGAAGACCTTTGTGTGTTG	1267
Db	1106	GTCTTAAAAAATTTCTTACGCTCATTTGCTGAAGAGAACCTTTGAGAGCTTTTAAAGTTG	1165
Qy	1268	CCAAAGTGCAGGTGGAAAATGGGAGCCAGGTGTTGGATGTCACACATGCGATGGCATGC	1327
Db	1166	CACAGCAACAGTGCAGAGCTGGCGCACAGATTTATGACATTTACATGATGAAGGATGC	1225
Qy	1328	TAGATGCTCAAGTGCAGAAATGACAGATTTTTCACCTTAAATTTGCTTCGAGCCAGACATCG	1387
Db	1226	TGATTCGCAAAATGCGATGCTGCAITTTCTTAAACCTTTGAGCATCCGAACTTGACATTT	1285
Qy	1388	CAAGGTTACCTTTGTGATCGACTCTCCCAATTTTCTGCTGATTTGAGCTGGGTAAAGT	1447
Db	1286	CACGTGTACCGATCATGATTTGACTCATCGAAATGGGAAATCATTTGAAGCTGGCTTAAAT	1345
Qy	1448	GCTGCCAAGGCAAGTGCATTTGCAATAGCATTTAGTCTGAAGGAGGAGAGAGACGACTTCT	1507
Db	1346	CGCTAAGATTAACCGTTGTTTAACTCAATTTCTTTAAAGAGGTTATGACGAGTTGCG	1405
Qy	1508	TGGAGAGGCGCAGGAAGATTAAGGATAGAGCTGCTATGTTGGTGCATGGCTTTTGATG	1567
Db	1406	TCGAAAAGCCGCTCTGCGCTCAATACGGTCTGCAATTAATGTTGATGCGCATTTGACG	1465
Qy	1568	AGGAGACAGGCAACAGAAACAGACACAAATTCAGAGTGTGCACCGCGCCCTACCATC	1627
Db	1466	AGTGTGTCAGGCGCACATCTGCTGAACGTAAACGTAAGCTCTGTAAGCGCTCTTACGACA	1525
Qy	1628	TGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTAATTTTGACCTTAATCTCTAA	1687
Db	1526	TTTTAGTAAAGAGTAGGCTTCCCTGCTGAGATATTAATTTTGACCGGACGCTGTTG	1585
Qy	1688	CCATTTGGATCGGAATGGAGAACACAACTTTGTATGTCATTAATTTTATCCATGCAACAA	1747
Db	1586	CAGTTGCAACTGGTATTGAAGAACACAACTACGCGATCGATTTTATTGAAGCAACGG	1645
Qy	1748	AGTCATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGCTTTTCCAACTTGTCTCT	1807
Db	1646	GCTGGATTAACAGAACTTACCGCACTTCCATGATTTCTGGTGGTGTGCTTAAAGCTTCGT	1705
Qy	1808	TCTCTTTCCGAGGAATGGAAGCCATTCGAGAAGCAATGCAATGGGTTTCTTCTTACCATG	1867
Db	1706	TCTCATTTCCGTTGGCATTAACAGTTCGTTGAGCCATTTCACTCTGTATTTCTGTACCATG	1765
Qy	1868	CAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGTATGATG	1927
Db	1766	CCATCAACCAAGCATGACCATGGTATTGTGAACGAGGTCAAAATGGCTATTATGATG	1825
Qy	1928	ATATCCATTAAGGAATCTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGACCTTGAGG	1987
Db	1826	ATATTCTTACCGACTAAAAGAGCGGTGAAGATGTCAATTTTAAATCAGAATCAAGGTG	1885
Qy	1988	CCACTGAGAGCTCTTACGTTTATGCCCAAGACTCAAGGCACA-----	2028
Db	1886	AGTCTGTGTAGGCTGCAACCGAGAAATTAATTTGAAGTTGCAAAAAATACCGTGGACAG	1945
Qy	2029	--GGAGGGAAGAAAGTCAATTCAGACTGATGAGTGGAGAAATTTGGCCCTGTGCAAGAACGCC	2086
Db	1946	GTGCTGCACAAAGAAAGCCGAGAACCTTGAATGGCGCAATGAATCAGTTTGAATAAAGCCTC	2005

Qy	2087	TTGAGTATGCCCTTGTGAAGGGCAATTGAAAAACAATATTATTGAGGATACCTGAGGAAGCCA	2144
Db	2086	TTGAATATGCTTGGTTTAAAGGTTATTACGACTTATTATTGACCAGACACCGAAGAAGCTC	2065
Qy	2147	GGTTAAACCAAAAAAATATATCCCGACCTCTCAATATAATTGAAGGACCCCTGATGAATG	2206
Db	2066	GCCTAAATCAAA-----CGTCTTTAGATGTGATGAAGGCCACTATAGACG	2116
Qy	2207	GAATGAAAAATTGTTGGTGATCTTTTTGGAGCTGGAAAAAATGTTTCTACCTCAGGTTATAA	2266
Db	2117	GCATGAAGCTCGTCGGTGACTTATTCGGTTTCAAGCAAAATGTTCTTGCCACAAGTTGTAA	2176
Qy	2267	AGTCAGCCCGGGTTATCAAGAGCGCTTGTGGCCACTTATCCCTTTTCATCGAAAAAGAAA	2326
Db	2177	AATCTGCCCGAGTCATGAACAAGACAGTGGCATGGCTCAACCCCGTACATCGAAGCTGAAA	2236
Qy	2327	GAGAAGAAACCAAGAGTCTTAAACGGCACAGTAGAAGAAAGAGGACCCCTTACACAGGACCA	2386
Db	2237	-----AGACCGAAGGACAGCTCTAAAGGTAAAG	2263
Qy	2387	TCGTCTGGCCCATGTTTAAAGGGCGACGTGCACGACATAGGCGAAGAACATAGTTGGAGTAG	2446
Db	2264	TCCTAATGGCAACGGTTAAAGGTGACGTACACGATATTGGTTAAAAATATTGTAGGCGTAG	2323
Qy	2447	TCCTTGGCTGCATAAATTTCCGAGTTATTGATTTAGGAGTCATGACTCCCATGTGATAAGA	2506
Db	2324	TACTTGGCTGTAAATGGCTATGACATTTGTGACCTTGGCGTAATGGTATACCTTTCGAGAGA	2383
Qy	2507	TACTGAAAGCTGCTCTTGACCAACAAAGCAGATATAATTGGCGCTGTACGAGCTCATCACTC	2566
Db	2384	TCCTACAACTGCAATTTGATGAAATAATGTGCATCATCGGATATTCTGGTCTGATCAACC	2443
Qy	2567	CTTCCCTGGATGAAATGATTTTTTGTGCCAAGGAAATGGAGAGATTAGCTATTAAGGATTC	2626
Db	2444	CATCTTTAGATGAAATGGTATTTGTCTTAAAGAAATGCAGCGTAAAGGCTTTAAACATTC	2503
Qy	2627	CATTGTTGATTGGAGGAGCAACCACTTCAAAAACCCACACAGCAGTTAAAATAG---CTC	2683
Db	2504	CTTTATTGATTGGTGGCGGCACTATTCTTAAAGCACATACAGCAGTAAAAAATTTGACCCCTC	2563
Qy	2684	CGAGATACAGTGCACCTGTAATCATGCTCTCGACCGCTCCAAAGAGTGTGCTGTGTGTT	2743
Db	2564	AGTATCAAAACGATCGGTAAATTTATGTTGCTGATGCTTCACGTGTGTTGGTGTAGCGA	2623
Qy	2744	CCGAGCTGTTAGATGAAAAATCTAAAGATGAATTAATTTGAGGAAATCATCGAAGAAATATG	2803
Db	2624	CAACCTTGCTTTCGAAAAGAAATGCTGGCGCATTTATTGAAAGAACATCGTGTGCAATATG	2683
Qy	2804	AAGATATTACACAGGACCAATTATGAGTCTCTCAAGGAGAGGAGATCTTACCCCTTAAGTC	2863
Db	2684	CCAAATTCGTGAGCGTTTAGCCAAACAAACAAACCAAGCGGCCCAACTGACTTTATAAG	2743
Qy	2864	AAGCAGAAAAAGTGGTTTCCAAATGGAATGGCTGTCTGAAACCTCCACCAAGTGAAGCCCA	2923
Db	2744	AGTCGGTTGAAATGGTTTTAAATTTGATGAAGCTACGTGCCAC-----CAAAACCAA	2797
Qy	2924	CGTTTATTGGGACCCAGGCTTTTGAAGCATATGACCTGCGAAGAGCTGGTGGACTACATTTG	2983
Db	2798	ATCTTTTGGGAAACACAGTTTTTAAAGAAATTTATCCGCTTTGCTACACTCGTGGATTTATTTT	2857
Qy	2984	ACTGGAGCCCTTTCTTGTGATGCTGGCAGCTCCGGGCAAGTACCCGAATCGAGGCTTCC	3043
Db	2858	ACTGGAGCCCACTCTTTATTTCTTTGGAGTTTAACTGGCAA-----ATTCC	2902
Qy	3044	CCAAGATATTTAACGACAAAAACAGTAGGTGGAGAGGCCAGGAAAGTCTACGATGATGCCCC	3103
Db	2903	CGAAAAATTTTAGAAGATGAAGTGGTCGGCGAAGCAGCAACTGACTTGTACAACCAAGCAC	2962
Qy	3104	ACAATATGCTGAACACACTGATTAGTCAAAAGAAAATCTCCGGGCCGGGGTGTGGTGGGT	3163
Db	2963	AAGCGATGTTGAAAGATATTATCGACAACAATCGTTTTGATGTCTGCTGCTGTATTATTTGTA	3022

RESIT.T 14

RESULTS 14  
US-09-489-039A-1418

US-09-409-039A-1418  
; Sequence 1418, Application US/09489039A

; Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 : TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: U

; CURRENT FILING DATE: 2000-01-27

;; PRIOR APPLICATION NUMBER: US 60

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 1418

; LENGTH: 3138

; TYPE: DNA

; ORGANISM:

;  
;  
FEATURE:  
1988/1989

NAME/KEY	LOCATION
;	
.	

; LOCATION: (3075)

; OTHER INFORMATION  
IIS-09-489-039A-1418

US-09-489-039A-1418



Query Match			17.9%; Score 701; DB 3; Length 3138;
Best Local Similarity			54.2%; Pred. No. 1.6e-203;
Matches 1648; Conservative			0; Mismatches 1311; Indels 81; Gaps 8;
Qy	149	GGATTATGGTCTGGATGGAGGGATGGGGACCATGATCCAGCGGAGAGCGTAACCGAAG	208
Db	146	GTAATTCGTGTTCTGGATGGGGGATGGGACCATGATCCAGGGCTATCGTCTGAGTGAGC	205
Qy	209	AACACTTCCGAGGTGAGGAATTAAGAATCATGTCAGCGCGCTCAAAAGGCAACAATGACA	268
Db	206	AGGACTTTCGCGGTGAGCGCTTTCGTGACTGGCGTGGACCTGAAGGCAACAACGACC	265
Qy	269	TTTTAAGTATAACTCAGCTGATGTCATTTTACCAATCCATAAGGAATPACTTGTGCTG	328
Db	266	TGCTGGTGCCTCAGCAAAACCGAAGTGATCCGCGAGATCCAGACGCCCTATTTCGAGGCG	325
Qy	329	GGGCGAGATATCATTTGAACAAATACTTTTAGCAGCACTAGTATTGGCCAAAGCTGACTATG	388
Db	326	CGCGGGATATCATTTGAGACCAACACTTTTAACTCGACGACCATCGCGATGCGGATTAAC	385
Qy	389	GCCTTGBACCTTGCCCTACCGGATGACATGTGCTCTGCAGGAGTGCCAG---AAAG	445
Db	386	AGATGGAATCCTGTTCGCGCAGATCAACTTTGCGCGCGGAAGCTGGCCCGCCAGCG	445
Qy	446	CTGCCGAGGAGTAACCTCCAGACAGGAATTAAGAGGTTTGTGGCAGGGGCTCTGGGTC	505
Db	446	CCGAGCCTGACCGGCCGCGACGCGGNAAAACCGCGTATGTGCGGGCGTCTTGCC	505
Qy	506	CGACTAATAAGCACTCTCTGTGCCCATCTGTGGAAAGGCGCGAATTAAGGAACATCA	565
Db	506	CGACCAACCGCACCGCTTCTATTTTCGCTGATGTGAACGATCCGCGCTTCGCTAATTA	565
Qy	566	CATTGTGATGCTTTGGAACATACCAAGACGAGGCCAAAGGACTTCTGATGCGCGGG	625
Db	566	CCTTTGACCACTGGTGGCGCCCTATCCGGAATCCACCGTGGCTGGTGGAAAGCGGCG	625
Qy	626	TTGATATCTTACTCATTTGAAACTATTTTGTGATCTGCCAATGCCAAGGACGCTTGTG	685
Db	626	TGGATCTGATCTGNTTGAACGGTGTGTTGATCTCTGAAAGCTTAAGGCGCCATCTATG	685
Qy	686	CACCTCAAAATCTTTTGGAGGAAATATGCTCCCGGCCCTATCTTTATTTTCAGGGACGA	745
Db	686	CGGTGAAGAGAGCTGGAGCGTTAGCGCTTGACCTGCCGCTGATGATCTCCGGACCA	745
Qy	746	TCGTTGATAAAGTGGCGGACTCTTCCGACAGACGAGGAGGGAATTTGTCATCAGCG	805
Db	746	TTACGATGCTCTGGCCGCGACGCTTCCGACAGACCAACCGAAGCCTTTTATACTCGC	805
Qy	806	TGTCCTCAGGAGAACCACTCTGCATTGGATTAATTTGTCTTTGGGTGCAGCTGAGATGA	865
Db	806	TGCGTCAAGCGGAGCCTTATCGTTTGGTCTGAACTGCGCCCTCGGGCGGACGAGCTGC	865
Qy	866	GACCTTTTATTGAATAATTTGGAAATGTACAAAGCCTATGTCTCTGTTATCCCAATG	925
Db	866	GGCAGTACGTCAGGAGCTGTCGCGCATTTGCGGAATGCTAGCTACCCGCGCACCGGAACG	925
Qy	926	CAGGTCTTCCCAACACTTTGGTGTGATGATAAAGCGCTTCTATGATGCGCAAGCACC	985
Db	926	CGGGCTGCCCAACCGCTTCGCGGAATACGATCTGGACGCCGACACCATGGCGAGCAAA	985
Qy	986	TAAAGGATTTTGTATGATGCTTGGTCAATATAGTTGGAGGATGCTGTGGTCAACAC	1045
Db	986	TTGCGGAATGGCGGAACCGGTTTCCCTCAATATCGTTGGCGGTGCTTGGGTACCAACC	1045
Qy	1046	CAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAAATTTGTAAGCCTTAGAGTTCCACCTG	1105
Db	1046	CGGAGCATATCGCCGCCATGAGCCGGCGGTAGCGGGCTGCGCGCGCTCAGTTGCGCG	1105
Qy	1106	CCACTGCTTTTGAAGGACATATGTTACTGTCTGGTCTAGAGCCCTTCAGGATTGGACCGT	1165
Db	1106	AAATTCGCGTGGCCTGCGCC-----CTCGCGCGCTGGAGCCGTTGAATATCGCGGACG	1159
Qy	1166	ACACCAACTTTGTTAACTTGGAGAGCGCTGTAATGTTTGCAGGATCAAGGAAGTTTGCTA	1225
Db	1160	ATAGCCTGTTTCGTTAACTCGTGGTGAACGAACATAACGTCACCGGCTCGCGGAAATTTAAGC	1219
Qy	1226	AACCTCATATGGCAGGAAACTATGAAGAAGCCTTGTGTGTTGCCAAGTGCAGGTGGAAA	1285
Db	1220	GGCTGATTAAAGAAGAGAAGTACAGGAAGCGCTGGACGTTGGCGCTCAGCAGGTGGAAA	1279
Qy	1286	TGGAGGCCAGGTGTTGGATGTCAACATATGGATGATGCATGCTAGATGGTCCAAAGTGCAA	1345
Db	1280	CGGCGCGCAGATTATCGATATCAACATGACGAGGGGATGCTCGACGCGGAACGCGGA	1339
Qy	1346	TGACCAATTTTGGCACTTAAATTTGCTTCCAGGCGACAGATCGCAAGGATACCTTTGTGCA	1405
Db	1340	TGGTGCCTTTCCTTAACTGATTTGCCGCGAGCGGACATTTGCCGCGTGCATCATGA	1399
Qy	1406	TGCACCTCTCCAAATTTTGTCTGATTTGAAGCTGGTTAAAGTGTCTGCAAGGGAAGTGCA	1465
Db	1400	TTGACTCCTCAAGTGGGAAGTCAATGAAGAAGGCCCTGAATAATGATTATTCAGGGCAAAAGCA	1459
Qy	1466	TTGTCAATAGCATTTAGTCTGAAGGAAGGAGGACGACTTCTTGGAGAAGGCCAGGAAGA	1525
Db	1460	TCGTTAACTCGATTTTCGATGAAGAGGGCGTCGAAATCCTTTATCCACCATGCGAACTGG	1519
Qy	1526	TTAAAAAGTATGAGCTGCTTATGTTGTCATGCTTTTGTGATGAAGAAAGGACAGGCAACAG	1585
Db	1520	TTGCGCGCTACGGCGCGCGTGGTGGTGTGATGCTTTTCGACGAGGTTGGTCAGGCCGACA	1579
Qy	1586	AAACAGACACAAAATCAGAGTGTGCACCCGGGCTTACCATCTGCTTGTGAAGAAAATCGG	1645
Db	1580	CCGCGAAGCTAAATTTGAGATTTTGGCGTCCGCGCTACAAAATCCTCACCGAAGAGTTCG	1639
Qy	1646	GCTTTAATCCAAATGACATTTATTTTGACCTTAATATCTTAACTTAACTTGGGACTCGAATGG	1705
Db	1640	GCTTCCCGCGGAGATATTATCTTCGACCCCAATATCTTTGCGCTGCCACCGGTATCG	1699
Qy	1706	AGGAACACAACTTGTATGCCATTTAATTTATCCATGCAACAAAGTCAATTAAGAAACAT	1765
Db	1700	AAGAGCATACAACCTACGCAAGGACTTTATTCGGCGCTCGGAAGACATCAACCGCGAGC	1759
Qy	1766	TACCTGGAGCCAGAAATTAAGTGGAGGTCTTTTCCAACTTGTCTCTCTTCCGAGGAATGG	1825
Db	1760	TGCGCGATGCGCTAATCTCCGCGCGGGTCTCCAAAGCTCTCTCTCTCTCCGCGGTATG	1819
Qy	1826	AAGCATTTCGAGAGCAATGATGCGGGTTCCTTTTACCAATCAATCAAGTCTCGCATGG	1885
Db	1820	ACCGGCTAGTGAGCGCATCCACGCGGTATTCCTCTACTAGCGCATCGCAACCGCATGG	1879
Qy	1886	ACATGGAGATAGTGAATGCTGGAAACCTCCCTGTTGATGATGATATCATTAAGGAACCTTC	1945
Db	1880	ACATGGCGATCGTCAACGCGGTCAAGCTGGCCATCTATGATGACCTCCCGGGCGAGCTGC	1939
Qy	1946	TGCAGCTCTGTGAAGATCTCATCTGGAATAAGACCTTAGGCCACTGAGGAGCTCTTAC	2005
Db	1940	GCGACGCGGTGGAAGACGTGATCTCAACCGCCGCGATGACAGTACCGAACGTTACTTG	1999
Qy	2006	GTTATGCCAGACTCAAGGCAACAGGAGGAAGAAGTCAATTCAGACTGAT-----G	2056
Db	2000	AGCTGGCGGAATAATATCGCGGAGCAAAAGCGGACGAGCGGCGAATGCCCAGCAGCGGG	2059
Qy	2057	AGTGAGAAATGGCCCTGTGGAAGAACGCTTTGATGATGCTTGTGAAGGGCAATTGAAA	2116
Db	2060	AGTGGCGCACCTGGGAAGTGAATAAAGCGCTCGAATATTCGCTGGTGAAGGGCAATTACCG	2119
Qy	2117	AACATATTTTGAAGGACTCTGAGGAGCGAGGTTAAACCAAAAAAATAATATCCCGACCTC	2176
Db	2120	AATTTTCGAACAGGACACCGAAGAGC-----GCGTCAGACAGCCGCCCGCCGA	2170
Qy	2177	TCAATATAATGAAGGACCCCTGATGAATGAATAAGAAATTTGTTGGTATCTTTTGGAG	2236
Db	2171	TTGAGGTCATTTGAAGGGCGCTGATGACGCGCATGAACGCTGTCGGCGATCTGTTGCGCG	2230
Qy	2237	CTGGAATAATGTTTCTACCTCAGGTTTAAAGTCAGGCCGCGGTTATGTAAGGAAGGCTGTTG	2296



Db 2231 AAGCAAAATGTTCTCTGCGCAGGTGGTGAATCGCCCGTGTATGTAAGCAAGCGCGTGG 2290  
Qy 2297 GCCACCTTATCCCTTTTCATGAAAGAAAGAGAGAAACAGAGTGTCTTAACCGGCACAG 2356  
Db 2291 CCTACCTTGAGCGGTTTATGAAGCCAGTA----- 2320  
Qy 2357 TAGAAGAGAGAGCCCTTACCAAGGCGCACCATCGTCTGCGCCACTGTGTTAAAGGCGACGTGC 2416  
Db 2321 ---AAGACGAGGCTCCAGTAAACGCAAAATGGTGAITGGCCACGGTAAAGGCGACGTTC 2377  
Qy 2417 ACACATAGGCAAGCAACATAGTTGAGTAGTCCCTTGGCTGCAATAATTTCCGAGTTATTG 2476  
Db 2378 ATGACATGGCAAAACATCGTCGCGGTGTCTGCACTGCATTAACATCAATCATCG 2437  
Qy 2477 ATTTAGGAGTCATGACTCCATGTGTAAGATACTGAAAGTGTCTTTGACCAACAAAGCAG 2536  
Db 2438 ATCTTGGCGTATGTTCCCGCGGATAAAATCTCNAACCGCCAAAGAGGTAAACGCGG 2497  
Qy 2537 ATATTAATTTGGCTGTTCAGACTCATCTCTTCCCTGGATGAATGATTTTGTGGCA 2596  
Db 2498 ATCTAAATTTGGTCTTTCCGGGCTGATTACTCCATCGCTGGACGATGGTTAATGTGGCA 2557  
Qy 2597 AGGAATCGAGAGATTACTCATAGGATTCCATTGTTGATTGGAGGCAACCACTTCAA 2656  
Db 2558 AAGAGATGGAGCGCCAGGCTTTACCATCCGCTGCTGATAGCGCGGACCACTCGA 2617  
Qy 2657 AAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACACTGTAAATCCATGTCTGG 2716  
Db 2618 AAGCGCATACGGGTGAATTCGAGCAGAACTACAGCGGCGCCGACCGTTACGTACAGA 2677  
Qy 2717 ACGGCTCCAGAGTGTGTGTGTTCCAGCTGTTAGATGAATAATTAAGAGATGAT 2776  
Db 2678 ACGCTCCGCGCACCGCTCGCGGTGTGTCTCGGCTGCTCTCGGACACGCGCGACGAGT 2737  
Qy 2777 ACTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGGACCATTTAGTCTCTCA 2836  
Db 2738 TTGTGCGCCGACGGCGAAGAGATGAACGGTGGTATCAGCATGGGCGTAAACAAAC 2797  
Qy 2837 AGAGAGAGAGATATCTTACCCTTTAGTCAAGCCAGAAAGTGGTTTCCAAATGATTTGGC 2896  
Db 2798 CGGCTACGCCCGCAGTGACCTTTCGCGCGCGGGGAAACGATCTGSCCTTTGACTGGG 2857  
Qy 2897 TGTCTGAACCTCACCGCTGAAGCCACGTTTATTTGGACCGAGTCTTTGAGACTATG 2956  
Db 2858 AGAGCTACACCGCGCGGTG---GCTCATCGTCTGGGAGTACAGCGGTGGAAGCCAGT 2913  
Qy 2957 ACCTGCAGAGCTGTGTGACTACATTTGACTGAGAGCTTTCTTTGATGCTCTGGCAGCTCC 3016  
Db 2914 A--TCGAGACTCTGCGCAACTACATCGACTGGACCGCGTCTTTCATGACCTGGTCTGG 2971  
Qy 3017 GGGCAAGTACCCGAATCGAGCTTCCCAAGATATTTAAGCAGCAAAACAGTAGTGGAG 3076  
Db 2972 CGGGTAATATCCGGC-----ATCTGGAAGATGAATGTAAGGAG 3016  
Qy 3077 AGGCAGGAAGTCTACGATGATGCCCAATATGCTGAACACACTGATTAAGTCAAAAGA 3136  
Db 3017 AGGCAGCAGCGCTGTTCAAAGACGCTAACGAGCTCTGGATAAGCTGAGCGCAGAGANA 3076  
Qy 3137 AACTCCGGGCGCGGGTGTGTGGGTTCTGGCCAGCACA 3176  
Db 3077 CCCTGAACCCCGCGCGGTGGTGGTCTGTGTTCCCGGGCCA 3116

## RESULT 15

US-09-252-991A-13893  
; Sequence 13893, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 13893  
; LENGTH: 2349  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13893

Query Match 15.7% Score 615.8; DB 3; Length 2349;  
Best Local Similarity 56.1%; Pred. No. 1.9e-177; Indels 21; Gaps 4;  
Matches 1251; Conservative 0; Mismatches 957

Qy 113 CCCTGCGGATGAGATCAATGCCCATTCTGCAGAAAGAGGATTATGGTCTGCTGATGGAGGGA 172  
Db 135 CCGCCTGCAAGCCCTTCAGCACGCCCTCAGGNAAGTATCTCTGATCTCTGATGGCGCA 194  
Qy 173 TGGGACCATGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTGAGGAATTTA 232  
Db 195 TGGSCACCATGATCCAGAGCTACAAGCTGGAAGAGGCCGACTACCGCGCGAGCGCTCG 254  
Qy 233 AAGATCATGCGAGCCCTGAAAGGCAACATGACATTTTAAGTATTAACCTCAGCTGATG 292  
Db 255 CCGACTGCGCGAGCGACGTGAAAGGCAACAAACGACCTCTTCTGCTGAGCGCGCGGACG 314  
Qy 293 TCATTTACCAATCCATAAGGATATCTTGTGGCTGGGCGAGATATATTGAACAAATA 352  
Db 315 TGATCCAGGCATCGAAGAGCCCTACTCGACGCGCGCGCGACATCTCTGAGACCAACA 374  
Qy 353 CTTTATGACGACCTAGTATTGCCCCAAGCTGACTATGSCCTTTGAACACTTTGSCCTTACC 412  
Db 375 CTTTCAAGCGCACCCAGGTGTCCAGGCCGACTACGCGCATGCGCTGCGCTACGAAAC 434  
Qy 413 TGAACATGTGCTCTGAGGAGTGGCCAGAAAGCTGCGAGGAGTAACTCTTCAGA--- 469  
Db 435 TCACCTCGAAGGGGCGCGCTGCGCCGCCAGCTGCGGAGCGCGAAGACCGCGCGAGACCC 494  
Qy 470 CAGGAATTAAGAGGTTTGTGCGAGGGCTCTGCGTCCGACTTAATAAGACACTCTCTGTGT 529  
Db 495 CGGCAAGCGCGCTTCTGCGCGGCTGTCTGCGCCCGACACGCGCGCACCTCTGCTGATTT 554  
Qy 530 CCCCATCTGTGGAAGGCCGGAATTAAGAACATCACTTATGATGAGCTTGTGGAAGCAT 589  
Db 555 CCGCGGACGTGAACAACCCCGCTTACCGCAAGCTCACTCTCGAGCAACTGCTGAGAACT 614  
Qy 590 ACCAAGACGAGGCCAAAGGACTTCTGATGCGGGGTGTGATATCTTACTCATTTGAAACTA 649  
Db 615 ACCTCAGGCGGACCCGCGGCTGATCGAAGCGCGCGCGACCTGATCTGATCGAGACCA 674  
Qy 650 TTTTGTATCTGCCAATGCAAGGACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 709  
Db 675 TCTTCGACACCTTCAACGCAAGCGCGCGCTTCTGCGCGCTCCAGGGCGGTGTTTCAGGAAC 734  
Qy 710 AATATGCTCCCGCGCTTCTTATTTTTCAGGAGCATGCTGTGATAAAGTGGGCGGACTC 769  
Db 735 TCGCGTGGAGCTGCCGATCATGATCTCCGGAACCATCAACGAGCGCTCCGCGCGGACCC 794  
Qy 770 TTTCCGACACAGAGGAGGATTTGTCTATCAGCGTGTCTCATGGAAGAACCACTCTGCA 829  
Db 795 TGTGCGGCGAGACCAACGAGGCTTCTGGAACCTCGGTGCGGCAATCGCCGCGGATCTCGG 854  
Qy 830 TTGGAATTAATTTGCTTTGGTGGTGCAGCTGAGATGAGACCTTTTATTAATTAATTTGAA 889  
Db 855 TAGGCTTGAATGCGCCCTCGCGCGCAAGGAATTTGCGCGCGGTATCATCAGGAACCTGCGA 914  
Qy 890 AATGTACAAACGCTATGCTCTCTGTTATCCCAATGAGGCTTCTCCCAACACCTTTGGTG 949  
Db 915 CCAAGGCGACACTCATGCTCTGCGGCCCAACCAAGCGCGGCTCCGCAACGCTTTCGCG 974  
Qy 950 ACTATGATGAACGCGCTTCTATGATGCGCAAGCACTTAAAGGATTTTCTATGATGCGCT 1009

Db 975 AATACGCAATCGCGCGAAATGGCGGTGGTTCGAGGAATTCGCGCGCGCGCT 1034  
Qy 1010 TGGTCAATATAGTTGGAGGATGCTGTGGTCAACACAGATCATATCAGGAAATTCGTG 1069  
Db 1035 TCCTCAATATCGTCGGCGGCTGCTGCGGCACACCCCGCGCACATCGAGGCGATCGCCA 1094  
Qy 1070 AAGCTGTGAATAATTTGAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGT 1129  
Db 1095 AGGAGTGGCCAGTACCCGCGCGGGCCATCCCGGAGATTCGCCGGCGCTGTC----- 1148  
Qy 1130 TACTGTCTGCTAGAGCCCTTCAGGATGAGGATGACCGATACACCAACTTTGTTAACTATGGAG 1189  
Db 1149 GCCTGTCCGCGCTGGAGCGGTTCCACATCGACCGCAGCTGCTGTTGTCAACTCGGCG 1208  
Qy 1190 AGCCTGTATGTTGAGGATCAAGGAAGTTTGTAACCTCATCATGCCAGGAACCTATG 1249  
Db 1209 AGCGACCAACATCACCGGTTTCGCGCAAGTTTCGCGCGCTGATCCGCGAGGAAACTACG 1268  
Qy 1250 AAGAAGCCTTGTGTTGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGATGTCA 1309  
Db 1269 CGGAAGCTCTCGAGGTGCGCCAGCAGCAGGTGGAGCGCGCCAGGTGATCGACATCA 1328  
Qy 1310 ACATGGATGATGCGATGCTAGATGTTCCAAAGTGCATGACAGATTTTGCACCTTAATTG 1369  
Db 1329 ACATGGACGAAGGCATGCTGACCTCGAAGCGGCGCATGGTCACCTTCCTCAACTGATCG 1388  
Qy 1370 CTTCCGAGCCAGACATCGCAAGGTACCTTTGTGATCGACTCTCTCCAACTTTTCTGTGA 1429  
Db 1389 CTTCCGAGCCGACATCTCGCGCGTGGCGATCATGTCGACTCTCTCCAACTGGAAGTGA 1448  
Qy 1430 TTGAAGCTGGGTAAAGTCTGCAAGGAAAGTGCATTTGCAATAGCATTTAGTCTGAAGG 1489  
Db 1449 TCGAGCGCGGCTGAAGTGATTCAGGCAAGGCGATCGTCAACTCGATTCGATGAAGG 1508  
Qy 1490 AAGGAGAGCGACTTCTTGGAGAGCGCAGGAAGATTTAAAGATGAGGACTGCTATGG 1549  
Db 1509 AAGGCGTCGAGGCTTCAAGCACCATGCCGCTGTGCAAGCGCTACGGCGCGGGTGG 1568  
Qy 1550 TGGTCATGGCTTTTGAAGAAGGACAGGACAGCAAGAAACAGACACAAAATCAGAGTGT 1609  
Db 1569 TGGTATGGCTTTCGACGAGGACGCGCACCCAGGCGCGCAAGGAAGAAATCT 1628  
Qy 1610 GCACCGGCGCTACCATCTGTTGAAAACCTGGGCTTTAATCCAAATGACATTTT 1669  
Db 1629 GCAAGCGCTCTCAGACATCTGTCGAGAGTGGCTTCCACCGGAGACATCATCT 1688  
Qy 1670 TTGACCTTAATATCTTAACCATTTGGGACTGGAATGGAGGAACACAACCTTTGATGCCATTA 1729  
Db 1689 TCGATGCGAAATCTTCGCCATCGCCACCGGCATCGAGGAACACAACACTACGGGTG 1748  
Qy 1730 ATTTTATCCATGCAACAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAG 1789  
Db 1749 ATTTTCATCAACGCTGCGCCTACATCCGCGACCACTGCGCTAGCGCTGAGCTCGGGCG 1808  
Qy 1790 GTCCTTCCAACTGCTCTCTTCGAGAGATGGAAGCCATTCGAGAAGCAATGCATG 1849  
Db 1809 GGGTGTCCAACTGCTCTCTTCGTCGCGGGAACACCCCGTGGCGGAGGATCCACT 1868  
Qy 1850 GGGTGTCTCTTACCATCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAA 1909  
Db 1869 CGGTGTCTCTTACTACGCGATCGCAACGCGCTGACCATGGGCATCGTCAACCGCGCC 1928  
Qy 1910 ACCTCCCTGTGTATGATGATTCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCT 1969  
Db 1929 AGCTGGAATCTTACGACGAGATTCGAAAGCGCTGCGGACCGGGTCCGAGACGTGTGC 1988  
Qy 1970 GGAATAAGACCTCGAGGCGCACTGAGAAGCTCTTACGTTATGC---CCAGACTCAAGGCA 2026  
Db 1989 TCAACCGCAGCGCCGAGGCGCCAGGCGCTGCTGGCGATCGCGACGACTACAGGGCG 2048  
Qy 2027 CAGGAGGGAAGAAAGTCAATTCAGACTGATGATGGAGAAATGGCCCTGTTCGAAGAACGCC 2086

Db 2049 GCAGCGCGGTCAAGGAGCGGAGGACGAGGAATGGCGCACTACAGTGTGAGAGCGCC 2108  
Qy 2087 TTGAGTATGCCCTTGTGAAGGGCATTTGAAAACATATTTATGAGGATACCTGAGGAAGCCA 2146  
Db 2109 TCAGAGCATGCGCTGGTCAAGGGCATCACCACTGGATCGTGGAGACACCGAGGA----- 2163  
Qy 2147 GGTAAACCAAAAAAATATCCCGACCTCTCAATATAAATTGAAGGACCCCTCATGAATG 2206  
Db 2164 -----ATCCCGCCAGCAGTGTGCGGTCCCATCGAGGTCAATCGAAGGTCCGCTGATGTCG 2219  
Qy 2207 GAATGAAAATTTGTTGGTGTATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAA 2266  
Db 2220 GGATGAACGTGGTGGCGACCTGTTGGCGCCGCGCAAGATGTTCTCCCGCAGGTGTCA 2279  
Qy 2267 AGTCAGCCCGGTTATGAAGAGGCTGTTGGCCACCTTATCCCTTTCATCGAAAAAGAA 2326  
Db 2280 AGTCCGCGCAGTGTGAACAGGCGGTGGCCACCTGATTCCTTTCATCGAGCGGAGA 2339  
Qy 2327 GAGAAAGAA 2335  
Db 2340 AAGCGATA 2348

Search completed: March 6, 2006, 19:02:39  
Job time : 494 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 18:54:41 / Search time 1963 Seconds  
(without alignments)  
16509.288 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcactgtggagacg.....ctcaaggaaatacaacctag 3919

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searches: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA Main:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3919	100.0	3919	7	US-10-607-712-1
2	3851	98.3	3856	7	US-10-607-712-75
3	914.6	23.3	3786	7	US-10-282-122A-25522
4	911.8	23.3	3684	7	US-10-282-122A-8842
5	902.6	23.0	3669	7	US-10-282-122A-32592
6	901.4	23.0	3705	7	US-10-282-122A-30149
7	899	22.9	3681	7	US-10-282-122A-19987
8	894	22.8	3705	7	US-10-282-122A-32118
9	892.2	22.8	3717	7	US-10-282-122A-33284
10	883.8	22.6	3684	7	US-10-282-122A-20596
11	883.2	22.5	3681	7	US-10-282-122A-40847
12	867.2	22.1	3684	7	US-10-282-122A-39570
13	859	21.9	3681	7	US-10-282-122A-24011
14	857.6	21.9	3771	7	US-10-282-122A-38912
15	854.4	21.8	3696	7	US-10-282-122A-41615
16	745.6	19.0	2796	7	US-10-282-122A-12513
17	708.6	18.1	3157	7	US-10-282-122A-36766
18	477.8	12.2	1779	7	US-10-282-122A-14755
19	416.8	10.6	420	3	US-09-796-692-8171
20	416.8	10.6	420	5	US-10-040-862-8171
21	416.8	10.6	420	6	US-10-057-475B-8171
22	416.8	10.6	420	8	US-10-154-884B-8171
23	416.8	10.6	420	8	US-10-764-324-8171

24	291.2	7.4	2418	3	US-09-974-300-220	Sequence 220, App
25	272.2	6.9	972	6	US-10-369-493-38170	Sequence 38170, A
26	272.2	6.9	972	6	US-10-369-493-38555	Sequence 38555, A
27	271.4	6.9	305	5	US-10-040-739-642	Sequence 642, App
28	268	6.8	1077	7	US-10-282-122A-13752	Sequence 13752, A
C 30	245.6	6.3	1065	7	US-10-282-122A-12010	Sequence 12010, A
C 31	245.6	6.3	1830121	8	US-10-329-670-1	Sequence 1, Appli
C 32	245.6	6.3	1830121	9	US-10-158-865-1	Sequence 1, Appli
C 33	235	6.0	1077	7	US-10-981-687-1	Sequence 1, Appli
34	224.6	5.7	861	6	US-10-282-122A-13844	Sequence 13844, A
35	205.6	5.2	1077	6	US-10-369-493-39757	Sequence 39757, A
36	199.2	5.1	3091	9	US-10-450-763-18370	Sequence 18370, A
37	198.2	5.1	2583	9	US-10-450-763-25658	Sequence 25658, A
38	198.2	5.1	2786	9	US-10-450-763-28834	Sequence 28834, A
C 40	183.8	4.7	3378	9	US-10-450-763-28667	Sequence 28667, A
41	183.8	4.7	3378	9	US-10-450-763-28792	Sequence 28792, A
42	183.8	4.7	3761	9	US-10-450-763-24468	Sequence 24468, A
43	183.8	4.7	4917	9	US-10-450-763-29901	Sequence 29901, A
44	141.4	3.6	901	5	US-10-027-632-121776	Sequence 121776, A
45	141.4	3.6	901	6	US-10-027-632-121776	Sequence 121776, A

ALIGNMENTS

RESULT 1  
US-10-607-712-1  
; Sequence 1, Application US/10607712  
; Publication No. US20040073018A1  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002005  
; CURRENT APPLICATION NUMBER: US/10/607,712  
; CURRENT FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 08/980,326  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/031,964  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: 60/050,310  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3919  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Other  
; LOCATION: (1)...(3919)  
; OTHER INFORMATION: Entire cloned cDNA encoding wild type methionine  
; OTHER INFORMATION: synthase.  
US-10-607-712-1

Query Match 100.0%; Score 3919; DB 7; Length 3919;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGTCACCTGTGGAGAGACGCTTCTTCTGCGCGCCCTCTGCGCAGGAGAGACTCGAC	60
Db	1	GGTCACCTGTGGAGAGACGCTTCTTCTGCGCGCCCTCTGCGCAGGAGAGACTCGAC	60
Qy	61	AACATGTCCACCGCGCTCCAGACCTGTGCGCAACCGAGCTGTGAGAAACCTCGCG	120
Db	61	AACATGTCCACCGCGCTCCAGACCTGTGCGCAACCGAGCTGTGAGAAACCTCGCG	120

121 GATGAGATCAATGCCATTCTCAGAGAGAGATTATGGTGCTGGATGAGGATGGGACC 180  
121 GATGAGATCAATGCCATTCTCAGAGAGAGATTATGGTGCTGGATGAGGATGGGACC 180  
181 ATGATCCAGCGGAGAGACTAAACGAAAGAACACTTCGAGGTCAGGAAATTTAAAGATCAT 240  
181 ATGATCCAGCGGAGAGACTAAACGAAAGAACACTTCGAGGTCAGGAAATTTAAAGATCAT 240  
241 GCCAGGCGCTGAAAGGCAACATGACATTTTAAAGTATTAACCTCAGCTGATGCTATTAC 300  
241 GCCAGGCGCTGAAAGGCAACATGACATTTTAAAGTATTAACCTCAGCTGATGCTATTAC 300  
301 CAATCCATTAAGGAATCTCTGCTGGCGGCGAGATATCATTTGAAGCAAAATPACTTTTAGC 360  
301 CAATCCATTAAGGAATCTCTGCTGGCGGCGAGATATCATTTGAAGCAAAATPACTTTTAGC 360  
361 AGCACTAGTATTTGCCAAGCTGACTATGCGCTTGAACACTTGGCTACCGGATGAACATG 420  
361 AGCACTAGTATTTGCCAAGCTGACTATGCGCTTGAACACTTGGCTACCGGATGAACATG 420  
421 TGCTCTGAGAGTGGCCAGAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAG 480  
421 TGCTCTGAGAGTGGCCAGAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAG 480  
481 AGGTTTGTGGCAGGGGCTCTGGGTCGAGCTTAATGAACACTCTCTGTGTCGCCATCTGTG 540  
481 AGGTTTGTGGCAGGGGCTCTGGGTCGAGCTTAATGAACACTCTCTGTGTCGCCATCTGTG 540  
541 GAAAGGCGGATATAGGAACATCACATTTGATGAGCTTTTGAAGCATACCAAGAGCAG 600  
541 GAAAGGCGGATATAGGAACATCACATTTGATGAGCTTTTGAAGCATACCAAGAGCAG 600  
601 GCCAAAGACCTTCTGGATGGCGGGTTGATATCTTAATCTCATTTGAAGCACTATTTTGAAT 660  
601 GCCAAAGACCTTCTGGATGGCGGGTTGATATCTTAATCTCATTTGAAGCACTATTTTGAAT 660  
661 GCCAATGCCAAGGAGCGCTTGTGCACTCCAAATCTTTTGAAGGAAATATGCTCCC 720  
661 GCCAATGCCAAGGAGCGCTTGTGCACTCCAAATCTTTTGAAGGAAATATGCTCCC 720  
721 CGGCTATCTTTATTTCAAGGACGATCGTGTATGATGAGGAGGAGGAGGAGGAGGAGGAG 780  
721 CGGCTATCTTTATTTCAAGGACGATCGTGTATGATGAGGAGGAGGAGGAGGAGGAGGAG 780  
781 ACAGAGAGGAGATTTGTCAATCAGCGTGTCTCATGAGGAGAACCACTCTGCAATTTGAAT 840  
781 ACAGAGAGGAGATTTGTCAATCAGCGTGTCTCATGAGGAGAACCACTCTGCAATTTGAAT 840  
841 TGTGCTTTGGGTCAGCTGAGATGAGACCTTTTATTTGAATATTTGGAATGATGACAA 900  
841 TGTGCTTTGGGTCAGCTGAGATGAGACCTTTTATTTGAATATTTGGAATGATGACAA 900  
901 GCCTATGCTCTGTATCCCAATGCAAGTCTTCCCAACACTTTTGGTGAATGATGAA 960  
901 GCCTATGCTCTGTATCCCAATGCAAGTCTTCCCAACACTTTTGGTGAATGATGAA 960  
961 ACGCTCTCTATGATGGCCAGCACTAAAGGATTTGCTATGATGAGGAGGAGGAGGAGGAG 1020  
961 ACGCTCTCTATGATGGCCAGCACTAAAGGATTTGCTATGATGAGGAGGAGGAGGAGGAG 1020  
1021 GTTGGAGATGCTGTGGGTCACACCATCATATCAGGGAATTTGCTGAGCTGTGAA 1080  
1021 GTTGGAGATGCTGTGGGTCACACCATCATATCAGGGAATTTGCTGAGCTGTGAA 1080  
1081 AATTGAAGCTTAGAGTTCCACCTGCCACTGTCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
1081 AATTGAAGCTTAGAGTTCCACCTGCCACTGTCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
1141 CTAGAGCCCTTCAGAGTTGAGCGGATACCAACTTTTGTGTTAACTTTGAGAGAGCGCTGTAAT 1200  
1141 CTAGAGCCCTTCAGAGTTGAGCGGATACCAACTTTTGTGTTAACTTTGAGAGAGCGCTGTAAT 1200

1201 GTTGCAGGATCAAGGAGATTTGCTAAACTCATCATGTCAGGAAACTATATGAGAGCCCTTG 1260  
1201 GTTGCAGGATCAAGGAGATTTGCTAAACTCATCATGTCAGGAAACTATATGAGAGCCCTTG 1260  
1261 TGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCAGGTGTTGGATGTCAACATGGATGAT 1320  
1261 TGTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCAGGTGTTGGATGTCAACATGGATGAT 1320  
1321 GGCATGCTAGATGGTCCAAGTGCAATGACCCAGATTTTGAACCTTAATTTGCTTCCGAGCCA 1380  
1321 GGCATGCTAGATGGTCCAAGTGCAATGACCCAGATTTTGAACCTTAATTTGCTTCCGAGCCA 1380  
1381 GACATCGCAAGGTACCTTTTGTGTCATCGACTCTCTCCAAATTTTGTGTGATTTGAAGCTGGG 1440  
1381 GACATCGCAAGGTACCTTTTGTGTCATCGACTCTCTCCAAATTTTGTGTGATTTGAAGCTGGG 1440  
1441 TTTAAAGTGTGCCAAGGAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAGGAGAGGAC 1500  
1441 TTTAAAGTGTGCCAAGGAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAGGAGAGGAC 1500  
1501 GACTTCTTGGAGAGGCCAGGAGATTTAAAGATTTGAGCTGCTATGTTGGTGGTCAATGGCT 1560  
1501 GACTTCTTGGAGAGGCCAGGAGATTTAAAGATTTGAGCTGCTATGTTGGTGGTCAATGGCT 1560  
1561 TTTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGCGCC 1620  
1561 TTTGATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGCGCC 1620  
1621 TACCATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAAT 1680  
1621 TACCATCTGCTTGTGAAAAAATCGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAAT 1680  
1681 ATCTTAACCATTTGGGACTGGAATGGAGGAAACAACTTGTATGTCATTAATTTTATCCAT 1740  
1681 ATCTTAACCATTTGGGACTGGAATGGAGGAAACAACTTGTATGTCATTAATTTTATCCAT 1740  
1741 GCAACAAAAGTCATTAAAGAAAACATTACTCTGGAGGACAGAAATAGTGAGAGTCTTTCCAA 1800  
1741 GCAACAAAAGTCATTAAAGAAAACATTACTCTGGAGGACAGAAATAGTGAGAGTCTTTCCAA 1800  
1801 TTGTCTTCTCTTCTCCGAGAAATGGAGGCAATTCGAGAGCAATGCAATGCGGTTTCTCTT 1860  
1801 TTGTCTTCTCTTCTCCGAGAAATGGAGGCAATTCGAGAGCAATGCAATGCGGTTTCTCTT 1860  
1861 TACCATGCAATCAAGTCTGCAATGGAGTGGAGTGGATGCTGGAACCTTCCCTG 1920  
1861 TACCATGCAATCAAGTCTGCAATGGAGTGGAGTGGATGCTGGAACCTTCCCTG 1920  
1921 TATGATGATATCCATAAGGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAC 1980  
1921 TATGATGATATCCATAAGGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATTAAGAC 1980  
1981 CCTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCAACAGGAGGAGGAGAA 2040  
1981 CCTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCAACAGGAGGAGGAGAA 2040  
2041 GTCATTCAGACTGATGATGGAGAAATGGGCTCTGTGAGAAACGCTTGTGATGATGCGCTT 2100  
2041 GTCATTCAGACTGATGATGGAGAAATGGGCTCTGTGAGAAACGCTTGTGATGATGCGCTT 2100  
2101 GTGAAGGCGATTTGAAAAACATATTTATTTGAGGATCTGAGGAGGAGGAGGAGGAGGAG 2160  
2101 GTGAAGGCGATTTGAAAAACATATTTATTTGAGGATCTGAGGAGGAGGAGGAGGAGGAG 2160  
2161 AAATATATCCCGACCTCTCAATATATATTAATTTGAAGGAGGAGGAGGAGGAGGAGGAG 2220  
2161 AAATATATCCCGACCTCTCAATATATTAATTTGAAGGAGGAGGAGGAGGAGGAGGAG 2220  
2221 GGTGATCTTTTGGAGCTGGAAGAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTT 2280  
2221 GGTGATCTTTTGGAGCTGGAAGAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTT 2280  
2281 ATGAAGAGGCTGTGGCCACCTTATCCCTTTTATGGAAGGAGGAGGAGGAGGAGGAGGAG 2340

```
Db 2281 ATGAGAAAGGCTGTGGCCACCTTATCCCTTTATGGAAGAAAGAGAGAAACAGAA 2340
Qy 2341 GTGCTTAAACGGCAGTAGAAGAGAGACCTTTACAGAGGACCACTCGTCTGGCCACT 2400
Db 2341 GTGCTTAAACGGCAGTAGAAGAGAGAGACCTTTACAGAGGACCACTCGTCTGGCCACT 2400
Qy 2401 GTTAAAGCCAGCTGACGACATAGGCAAGACATAGTTGGAGTAGTCTGGCTGCAAT 2460
Db 2401 GTTAAAGCCAGCTGACGACATAGGCAAGACATAGTTGGAGTAGTCTGGCTGCAAT 2460
Qy 2461 AATTTCCGAGTTATTTAGGATCATGACTCCATGATAGATACCTGCTGCTGCT 2520
Db 2461 AATTTCCGAGTTATTTAGGATCATGACTCCATGATAGATACCTGCTGCTGCT 2520
Qy 2521 CTTGACCAACAAGCAGATATAATTTGGCTCTGAGGACTCATCTCTCCCTGGATGAA 2580
Db 2521 CTTGACCAACAAGCAGATATAATTTGGCTCTGAGGACTCATCTCTCCCTGGATGAA 2580
Qy 2581 ATGATTTTGTGCAAGGAAATGGAGAGATAGCTATAGATTAAGATTCATTTGATGGA 2640
Db 2581 ATGATTTTGTGCAAGGAAATGGAGAGATAGCTATAGATTAAGATTCATTTGATGGA 2640
Qy 2641 GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCT 2700
Db 2641 GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCT 2700
Qy 2701 GTAATCCTATCTCTGAGCGCTCAAGAGTGTGTGTGTGTTCCACAGCTGTTAGATGAA 2760
Db 2701 GTAATCCTATCTCTGAGCGCTCAAGAGTGTGTGTGTGTTCCACAGCTGTTAGATGAA 2760
Qy 2761 AATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATAGACAGGAC 2820
Db 2761 AATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATAGACAGGAC 2820
Qy 2821 CATTTAGGCTCTCAAGGAGAGAGATATCTTACCTTTAGTCAAGCCAGCAAAAAGTGT 2880
Db 2821 CATTTAGGCTCTCAAGGAGAGAGATATCTTACCTTTAGTCAAGCCAGCAAAAAGTGT 2880
Qy 2881 TTCCAAATGATGGCTGTCTGAACCTCACCAAGTGAAGCCACGTTTATTTGGACCCAG 2940
Db 2881 TTCCAAATGATGGCTGTCTGAACCTCACCAAGTGAAGCCACGTTTATTTGGACCCAG 2940
Qy 2941 GTCTTTGAAGACTATGACCTGCAAGAGCTGGTGGACTACATGACTGGAAGCCTTTCTTT 3000
Db 2941 GTCTTTGAAGACTATGACCTGCAAGAGCTGGTGGACTACATGACTGGAAGCCTTTCTTT 3000
Qy 3001 GATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGCAG 3060
Db 3001 GATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGCAG 3060
Qy 3061 AAAACAGTAGTGAGAGGCGCAGGAAGTCTACGATGATGCCACCAATATGCTGAACACA 3120
Db 3061 AAAACAGTAGTGAGAGGCGCAGGAAGTCTACGATGATGCCACCAATATGCTGAACACA 3120
Qy 3121 CTGATTAGTCAAAAGAACTCCGGGCGGGGTGTGTTGGTTCGCGCCAGCAGACAGT 3180
Db 3121 CTGATTAGTCAAAAGAACTCCGGGCGGGGTGTGTTGGTTCGCGCCAGCAGACAGT 3180
Qy 3181 ATCCAAGACGATTAACCTGTACGAGAGGCTGTGTGCCCAAGGCTGACAGAGCCCATTA 3240
Db 3181 ATCCAAGACGATTAACCTGTACGAGAGGCTGTGTGCCCAAGGCTGACAGAGCCCATTA 3240
Qy 3241 GCCATTTCTATGGTTTAAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATAC 3300
Db 3241 GCCATTTCTATGGTTTAAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATAC 3300
Qy 3301 TACTGCCTCTCAGACTTTCATCGCTCCCTTGCAATCTCGCATCCGCTGACTACCTGGGCTG 3360
Db 3301 TACTGCCTCTCAGACTTTCATCGCTCCCTTGCAATCTCGCATCCGCTGACTACCTGGGCTG 3360
Qy 3361 TTGCGCTTGCCTGCTTTGGGTAGAAAGCTGAGCAAGGCTTATGAGGATGATGTGAC 3420
```

```
Db 3361 TTGCGCTTGCCTGCTTTGGGTAGAAAGCTGAGCAAGGCCCTATGAGGATGATGTGAC 3420
Qy 3421 GACTACAGCAGCATCATGTCAAGCGCTGGGGGACCGGCTGCGCAGAGCGCTTTGACAGAA 3480
Db 3421 GACTACAGCAGCATCATGTCAAGCGCTGGGGGACCGGCTGCGCAGAGCGCTTTGACAGAA 3480
Qy 3481 GAGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGAGCTGGAC 3540
Db 3481 GAGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGAGCTGGAC 3540
Qy 3541 GTGCGAGACCTCGCAAGGTTGGCGTCAAGGGGCATCCGCCCGGCTCTGCTACCCCGAGC 3600
Db 3541 GTGCGAGACCTCGCAAGGTTGGCGTCAAGGGGCATCCGCCCGGCTCTGCTACCCCGAGC 3600
Qy 3601 CAGCCCGACACACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACA 3660
Db 3601 CAGCCCGACACACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACA 3660
Qy 3661 GGCATTAGTTAACAGAACTTAGCAATGGCACTGCTTCCAGCAGTCTCAGGCTCTAC 3720
Db 3661 GGCATTAGTTAACAGAACTTAGCAATGGCACTGCTTCCAGCAGTCTCAGGCTCTAC 3720
Qy 3721 TTCTCCAAATTTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAAAGGATCAGGTTGAG 3780
Db 3721 TTCTCCAAATTTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAAAGGATCAGGTTGAG 3780
Qy 3781 GATTATGCAATTTGAGGAAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATT 3840
Db 3781 GATTATGCAATTTGAGGAAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATT 3840
Qy 3841 TTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTGGCTTTTATCTGATGATCC 3900
Db 3841 TTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTTTTTGGCTTTTATCTGATGATCC 3900
Qy 3901 TCAAGGAAATACAACTAG 3919
Db 3901 TCAAGGAAATACAACTAG 3919
```

## RESULT 2

```
US-10-607-712-75
; Sequence 75, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 3856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(3856)
; OTHER INFORMATION: nnn at positions 2640-2642 is either AAT or no
; OTHER INFORMATION: nucleotides; n at position 2756 is either A or G;
; OTHER INFORMATION: n at position 2758 is either C or G.
```

US-10-607-712-75

Query Match		98.3%;	Score 3851;	DB 7;	Length 3856;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 3851;		Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;	
Qy	64	ATGTCACCGCGCTCCAGACCTCTGCGCAACCCGGAAGTCTGGAAGAAACCCCTGCGGAT	123			
Db	1	ATGTCACCGCGCTCCAGACCTCTGCGCAACCCGGAAGTCTGGAAGAAACCCCTGCGGAT	60			
Qy	124	GAGATCAATGCAATCTGCGAAGAGGATTAATGCTGCGATGAGGATGAGGACCAATG	183			
Db	61	GAGATCAATGCAATCTGCGAAGAGGATTAATGCTGCGATGAGGATGAGGACCAATG	120			
Qy	184	ATCCAGCGGAGAGCTAAACGAAAGACATCTCCGAGTCAAGGATTAAGATCATGACC	243			
Db	121	ATCCAGCGGAGAGCTAAACGAAAGACATCTCCGAGTCAAGGATTAAGATCATGACC	180			
Qy	244	AGGCCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCAATTAACAA	303			
Db	181	AGGCCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGTCAATTAACAA	240			
Qy	304	ATCCATAAGGAATACTTGGCTGGGCGAGATATCATTTGAACAAATATCTTTAGCAGC	363			
Db	241	ATCCATAAGGAATACTTGGCTGGGCGAGATATCATTTGAACAAATATCTTTAGCAGC	300			
Qy	364	ACTAGTATGCCCAGCTGACTATGGCTTCAACACTTGGCTACCGGATGAACATGTC	423			
Db	301	ACTAGTATGCCCAGCTGACTATGGCTTCAACACTTGGCTACCGGATGAACATGTC	360			
Qy	424	TCTGAGAGTGGCCAGAAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGAGG	483			
Db	361	TCTGAGAGTGGCCAGAAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGAGG	420			
Qy	484	TTTGTGCGAGGGCTTGGGTCGACATTAAGACACTCTCTGTGTCCTCCATCTGTGAA	543			
Db	421	TTTGTGCGAGGGCTTGGGTCGACATTAAGACACTCTCTGTGTCCTCCATCTGTGAA	480			
Qy	544	AGGCCGATTAAGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACAGGCC	603			
Db	481	AGGCCGATTAAGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACAGGCC	540			
Qy	604	AAAGGACTCTGGATGGCGGGTGTATCTTACTCATTTGAACATTAATTTGATCTGCC	663			
Db	541	AAAGGACTCTGGATGGCGGGTGTATCTTACTCATTTGAACATTAATTTGATCTGCC	600			
Qy	664	AATGCCAAGGCGCTTGTGTCCTCAAAATCTTTTGAAGGAAATATGCTCCCGG	723			
Db	601	AATGCCAAGGCGCTTGTGTCCTCAAAATCTTTTGAAGGAAATATGCTCCCGG	660			
Qy	724	CCTATCTTTATTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCCGACAGACA	783			
Db	661	CCTATCTTTATTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCCGACAGACA	720			
Qy	784	GGAGAGGATTTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATTTGAATTAATGT	843			
Db	721	GGAGAGGATTTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATTTGAATTAATGT	780			
Qy	844	GCTTTGGGTGAGCTGAGATGAGACCTTTTATGAAATTAATTTGAAGGAAATGACACGCC	903			
Db	781	GCTTTGGGTGAGCTGAGATGAGACCTTTTATGAAATTAATTTGAAGGAAATGACACGCC	840			
Qy	904	TATGTCCTCTGTTATCCCAATGAGGCTCTCCCAACACTTTTGGTGATGATGATGAACG	963			
Db	841	TATGTCCTCTGTTATCCCAATGAGGCTCTCCCAACACTTTTGGTGATGATGATGAACG	900			
Qy	964	CCTTCTATGATGGCCCAAGCACTTAAGGATTTTGTCTATGGATGGCTTGGTCAATATAGTT	1023			
Db	901	CCTTCTATGATGGCCCAAGCACTTAAGGATTTTGTCTATGGATGGCTTGGTCAATATAGTT	960			
Qy	1024	GGAGGATGCTGTGGGTCAACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAT	1083			
Db	961	GGAGGATGCTGTGGGTCAACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAT	1020			

Qy	1084	TGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTCTA	1143
Db	1021	TGTAAGCCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTCTA	1080
Qy	1144	GAGCCCTTCAGGATGGACCGTACA CCACTTTTGTAAATTTGGAGAGCCGCTGTAATGTT	1203
Db	1081	GAGCCCTTCAGGATGGACCGTACA CCACTTTTGTAAATTTGGAGAGCCGCTGTAATGTT	1140
Qy	1204	GCAGGATCAAGGAAGTTTGTCTAACTCATCATGGCAGGAACTATGAGAGGCTTGTGT	1263
Db	1141	GCAGGATCAAGGAAGTTTGTCTAACTCATCATGGCAGGAACTATGAGAGGCTTGTGT	1200
Qy	1264	GTTCGCAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCACATGGAATGATGGC	1323
Db	1201	GTTCGCAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCACATGGAATGATGGC	1260
Qy	1324	ATGCTAGATGGTCCAAAGTGCATGACAGATTTTGCACATTAATTTGCTTCCGAGCCAGAC	1383
Db	1261	ATGCTAGATGGTCCAAAGTGCATGACAGATTTTGCACATTAATTTGCTTCCGAGCCAGAC	1320
Qy	1384	ATGCAAGGTACTTTTGTGTCATCGACTCTCCAAATTTTGTCTGTGATTGAGCTGGGTTA	1443
Db	1321	ATGCAAGGTACTTTTGTGTCATCGACTCTCCAAATTTTGTCTGTGATTGAGCTGGGTTA	1380
Qy	1444	AAGTGTGCGCAAGGAAGTGCAATTTGTCAAATAGCATTAGTCTGAAGGAAGGAGGACGAC	1503
Db	1381	AAGTGTGCGCAAGGAAGTGCAATTTGTCAAATAGCATTAGTCTGAAGGAAGGAGGACGAC	1440
Qy	1504	TTCTTGGAGAGGCCAGGAAGATTAAAAAGTATGGAGCTGCTATGTGGTGCATGGCTTTT	1563
Db	1441	TTCTTGGAGAGGCCAGGAAGATTAAAAAGTATGGAGCTGCTATGTGGTGCATGGCTTTT	1500
Qy	1564	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCTAC	1623
Db	1501	GATGAAGAGGACAGGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCTAC	1560
Qy	1624	CATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATC	1683
Db	1561	CATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATC	1620
Qy	1684	CTAACCATTTGGGACTGGAATGGAGGAA CACAACCTTGTATGCCATTAATTTTATCCATGCA	1743
Db	1621	CTAACCATTTGGGACTGGAATGGAGGAA CACAACCTTGTATGCCATTAATTTTATCCATGCA	1680
Qy	1744	ACAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATTAAGTGGAGCTCTTCCAACTTG	1803
Db	1681	ACAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATTAAGTGGAGCTCTTCCAACTTG	1740
Qy	1804	TCCTTCTCTTCCGAGGAATGGAAAGCCATTCGAGAAAGCAATGATGGGGTTTTCCTTTAC	1863
Db	1741	TCCTTCTCTTCCGAGGAATGGAAAGCCATTCGAGAAAGCAATGATGGGGTTTTCCTTTAC	1800
Qy	1864	CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTCGAAACCTCCCTGTGTAT	1923
Db	1801	CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTCGAAACCTCCCTGTGTAT	1860
Qy	1924	GATGATATCAATTAAGAACTTCTGAGCTCTGTGAGATCTCATCTGGAATTAAGACCCCT	1983
Db	1861	GATGATATCAATTAAGAACTTCTGAGCTCTGTGAGATCTCATCTGGAATTAAGACCCCT	1920
Qy	1984	GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCA CAGGAGGGAAGAAAGTC	2043
Db	1921	GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCA CAGGAGGGAAGAAAGTC	1980
Qy	2044	ATTGAGACTGATGAGTGGAAATGGCCCTGTGGAAGACGCTTGTAGTATGCCCCTTTGTG	2103
Db	1981	ATTGAGACTGATGAGTGGAAATGGCCCTGTGGAAGACGCTTGTAGTATGCCCCTTTGTG	2040
Qy	2104	AAGGCCATTGAAAAACATATTTATTGAGGATATCTGAGGAAGCCAGGTTTAAACAAAAAAA	2163
Db	2041	AAGGCCATTGAAAAACATATTTATTGAGGATATCTGAGGAAGCCAGGTTTAAACAAAAAAA	2100



Qy	2164	TATCCCGAGCTCTCAATATATAATGGAAGGACCCCTGTAGTAATGGAATGGAATGGAATGTTGGT	2222
Db	2101	TATCCCGAGCTCTCAATATATAATGGAAGGACCCCTGTAGTAATGGAATGGAATGTTGGT	2160
Qy	2224	GATCTTTTTCGAGCTCGAAATAATGTTTCTACCTCAGGTATAAAGTCAGCCCGGGTTATG	2283
Db	2161	GATCTTTTTCGAGCTCGAAATAATGTTTCTACCTCAGGTATAAAGTCAGCCCGGGTTATG	2220
Qy	2284	AAGAAGGCTGTGTGGCCACCTTATCCCTTTTCATGSAAAAAAGAGAAGAAACCGAGATG	2343
Db	2221	AAGAAGGCTGTGTGGCCACCTTATCCCTTTTCATGSAAAAAAGAGAAGAAACCGAGATG	2280
Qy	2344	CTTTAAACGGCAGCTAGTAAGAAGAGGACCCCTTACCGGGCCACCATCGTCTGGGCCACTGTT	2403
Db	2281	CTTTAAACGGCAGCTAGTAAGAAGAGGACCCCTTACCGGGCCACCATCGTCTGGGCCACTGTT	2340
Qy	2404	AAAGGGCAGCTGCACGACATATAGGCAAGAACATAGTTGGAGTGTAGTCCCTTGGCTGCAATAAT	2463
Db	2341	AAAGGGCAGCTGCACGACATATAGGCAAGAACATAGTTGGAGTGTAGTCCCTTGGCTGCAATAAT	2400
Qy	2464	TTCCGAGTTATTGATTTTAGGAGTCATGACTCCATGTGATTAAGTACTTGAAGCTGCTCTTT	2523
Db	2401	TTCCGAGTTATTGATTTTAGGAGTCATGACTCCATGTGATTAAGTACTTGAAGCTGCTCTTT	2460
Qy	2524	GACCAAAAGCAGATATAATTGGCGCTGTCAAGACTCATCACTCTTCCCTGGATGAATAATG	2583
Db	2461	GACCAAAAGCAGATATAATTGGCGCTGTCAAGACTCATCACTCTTCCCTGGATGAATAATG	2520
Qy	2584	ATTTTGTTCGCAAGGAAATGGAGAGATTAGCTATTAAGGATTCCTATGTTGATGAGGGA	2643
Db	2521	ATTTTGTTCGCAAGGAAATGGAGAGATTAGCTATTAAGGATTCCTATGTTGATGAGGGA	2580
Qy	2644	GCAACCACTTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTCACCTGTA	2703
Db	2581	GCAACCACTTTCAAAAAACCCACACAGCAGTTTAAAAATAGCTCCGAGATACAGTCACCTGTA	2640
Qy	2704	ATCATGTCTGTGACGCGTCCAAAGAGTGTGGTGTGTTCCTCAGCTGTTTAGATGAATAAT	2763
Db	2641	NNCCATGTCTGTGACGCGTCCAAAGAGTGTGGTGTGTTCCTCAGCTGTTTAGATGAATAAT	2700
Qy	2764	CTAAAGGNTGAATACTTTTGAGGAAATCATGGAAGNATATGAAGATATTAGACAGGACCAT	2823
Db	2701	CTAAAGGNTGAATACTTTTGAGGAAATCATGGAAGNATATGAAGATATTAGACAGGNCAT	2760
Qy	2824	TATGAGTCTCTCAAGGAGAGGAGATCTTACCCTTAAGTCAAGCCAGCAAAAAAGTGGTTTC	2883
Db	2761	TATGAGTCTCTCAAGGAGAGGAGATCTTACCCTTAAGTCAAGCCAGCAAAAAAGTGGTTTC	2820
Qy	2884	CAAAATGAATTTGGTGTCTGAACTCACCAGTGAAGCCCACTTTATTTGGGACCCAGGTC	2943
Db	2821	CAAAATGAATTTGGTGTCTGAACTCACCAGTGAAGCCCACTTTATTTGGGACCCAGGTC	2880
Qy	2944	TTTTGAAGACTATGACCTGCGAAGAGCTGGTGGACTACATTCAGCTGGAAGCCCTTTCTTGAT	3003
Db	2881	TTTTGAAGACTATGACCTGCGAAGAGCTGGTGGACTACATTCAGCTGGAAGCCCTTTCTTGAT	2940
Qy	3004	GTCTGGCAGCTCCGGGGCAAGTACCGAAATCGAGGCTTTCCCCCAAGATATTAAACGACAAA	3063
Db	2941	GTCTGGCAGCTCCGGGGCAAGTACCGAAATCGAGGCTTTCCCCCAAGATATTAAACGACAAA	3000
Qy	3064	ACAGTAGGTGGAGGCGCAGGAAGGTCTACGATGATGCCCAACAATATGCTGGAACACACTG	3123
Db	3001	ACAGTAGGTGGAGGCGCAGGAAGGTCTACGATGATGCCCAACAATATGCTGGAACACACTG	3060
Qy	3124	ATTAGTCAAAAGAAACTTCGGGGCCCGGGGTGTGGTTGGGTTCGTGGCCAGCACAGATATC	3183
Db	3061	ATTAGTCAAAAGAAACTTCGGGGCCCGGGGTGTGGTTGGGTTCGTGGCCAGCACAGATATC	3120
Qy	3184	CAGACGACATTCACCTGTACGAGAGGCTGCTGTGCCCAAGGCTCGAGAGCCCATAGCC	3243
Db	3121	CAGACGACATTCACCTGTACGAGAGGCTGCTGTGCCCAAGGCTCGAGAGCCCATAGCC	3180
Qy	3244	ACTTTCTATGGGTTTAAGGCAACAGGCTGAGAAAGACTCTGCCAGCAGCGAGCCATACTAC	3303

Db	3181		ACTTCTCATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCATCGGAGCCATACTAC	3240
Qy	3304	TGCCTCTCAGACTTCATCGCTCCCTTGTCATCTCGGCATCCGTGACTACCTGGGCCCTGTTT	3363	
Db	3241	TGCCTCTCAGACTTCATCGCTCCCTTGTCATCTCGGCATCCGTGACTACCTGGGCCCTGTTT	3300	
Qy	3364	GCGTTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTAGACAC	3423	
Db	3301	GCGTTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTAGACAC	3360	
Qy	3424	TACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGCGCAGAGGCCCTTTCAGAGAAG	3483	
Db	3361	TACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGCGCAGAGGCCCTTTCAGAGAAG	3420	
Qy	3484	CTCCATGAAGAAGTTCGCCCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGCAGTTCGACGTC	3543	
Db	3421	CTCCATGAAGAAGTTCGCCCGAGAACTGTGGGCCCTACTGTGGCAGTGAAGCAGTTCGACGTC	3480	
Qy	3544	GCAGACTGCGAAGTTGGGGTACAAGGGCATCCGCCCGGCTCTTGGTACCCAGCCAC	3603	
Db	3481	GCAGACTGCGAAGTTGGGGTACAAGGGCATCCGCCCGGCTCTTGGTACCCAGCCAC	3540	
Qy	3604	CCGCACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGACGAGTCTACAGGC	3663	
Db	3541	CCGCACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGACGAGTCTACAGGC	3600	
Qy	3664	ATTAGGTTTAACAGAACTATTAGCAATGGCACCTGTCTTACGACAGTCTCAGGCCCTACTTTC	3723	
Db	3601	ATTAGGTTTAACAGAACTATTAGCAATGGCACCTGTCTTACGACAGTCTCAGGCCCTACTTTC	3660	
Qy	3724	TCCAAATTGAAGTCCAAATAATTTTGTGTGGGGAAGATTTCCAAGGATCAGGTTGAGGAT	3783	
Db	3661	TCCAAATTGAAGTCCAAATAATTTTGTGTGGGGAAGATTTCCAAGGATCAGGTTGAGGAT	3720	
Qy	3784	TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTCGACCCATTTC	3843	
Db	3721	TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTCGACCCATTTC	3780	
Qy	3844	GGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTGGCTTTTTTATCTTGGATGATCCTCA	3903	
Db	3781	GGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTGGCTTTTTTATCTTGGATGATCCTCA	3840	
Qy	3904	AGGAAATACAACTTAG	3919	
Db	3841	AGGAAATACAACTTAG	3856	

### RESULT. T 3

```

US-10-282-122A-2552
RESULI 3
Sequence 2552, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essene
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23

```



1813 TTCCAGCCATCAAGCGCGCTGACATGGGATCGTCAAGCGCGCGCTGGTGGCC 1872  
1921 TATGATGATATCCATAAGAACTTCTGAGCTCTGTGAAGATCTCATCTGGAATPAAGAC 1980  
1873 TACGACTCGATCGACCCCGAGCTCGGACCGCATCGAGGACGTCGTGCTGAACCGCGC 1932  
1981 CCTGAGGCCACTGAGAGCTCTTACGTTATGCCAG-----ACTCAAGGACAGGAGGG 2034  
1933 GAGGACCGCGCGAGAGGCTCTGGAGATCGCCGAGCGGTTCAACNAAGTCGCGAGATGCT 1992  
2035 AAGAAAGTCATTACAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCTTGAATAT 2094  
1993 TCGAGGATTCGGCGCGCGAGTGGCGCGCTGCGGTCCGCGAGCGGATCACGCAC 2052  
2095 GCCCTTGTGAAGGCGATGAAAAACATATATTAGAGATCTAGAGGAGCGAGG---TTA 2151  
2053 GCCCTGTGTCAAGGCGATCGACGCCAGCTCGAGGATCGAGGCGCGCTGTGAGCGCATG 2112  
2152 AACCAAAAAAATATCCCGACCTCTCAATATAATTGAAGACCCCTGATGAATGGAATG 2211  
2113 ATTCCCGCGCGGTGGCGCGCGATCGAGGTGATCGAGGCGCGCTGTGAGCGCATG 2172  
2212 AAAATTGTGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTTATAAGTCA 2271  
2173 AACGTGTGGGACCTGTTCGGCTCCGCGAGATGTTTTTCCCGCAGGTGGTGAATCG 2232  
2272 GCCCGGTTATGAAGAAGGCTGTGGCCACCTTATCCCTTTTCATGGAAAAAGAGAGAA 2331  
2233 GCCCGGTGTGAAGAAGGCGCTGCCCTACCTGCTGCCGTTTCATCGAGGCGGAGAGAA 2292  
2332 GAAACGAGGTCTTAACGCGACAGTAGTAGAAGAGACCTTACAGAGGACCATCGTG 2391  
2293 GAGTCCGCGGTGTCGGTTCCA-----AGGACACCAACGCGCACCATCGTG 2337  
2392 CTGSCCACTGTTAAAGCGAGCTGCAGCATAGCGCAAGACATAGTTGGAGTAGTCCTT 2451  
2338 ATGGCGACCGTCAAGGCGAGCTCACGACATCGCGAAGACATCTGTCGGAGTTGCTTG 2397  
2452 GGCTGCAATAATTCCGAGTTATTGATTTAGGAGTCAATGCTCATGTGATGAAGATCTG 2511  
2398 CAGTGCAACAACTTCGAGGTGATCGACCTCGGTGTGATGGTCCGCGCCAGAGATCTC 2457  
2512 AAAGTCTCTTTGACCAAGAGAGATATATTGGCTGTGAGAGCTCATCACTCTCTCC 2571  
2458 GATCTCGAAGAGACGACGCGCGACATCATCGGCTGTGCGGCTGATCACCCCGTCC 2517  
2572 CTGATGAAATGATTTTGTGCAAGAAATGAGAGATTAGCTATAAGGATTCATTTG 2631  
2518 CTGAGCAGAGATGCCAACTTCGCGCGCGAGATGGAACGCGAGGCGCTCGAGATCCCGCTG 2577  
2632 TTGATTGGAGGAGCAACCACTTCAAAAAACCCACAGCAGTTAAATAGCTCCGAGATAC 2691  
2578 CTGATCGCGGCGGACCACTTCGCGCGCCCAACAGCGCGTGAAGATCTCGCGCGCGGT 2637  
2692 AGTGCACCTGTATTCATGCTCGGACGCTGCAAGAGTGTGAGTGTGTTCCAGCTG 2751  
2638 TCCGCGCGGTGTGCTGGGTCAAGAGCGCTTCCCGCTCGGTGCGGTGCGCGCGCGCTG 2697  
2752 TTAGATCAAAATCTAAGGATGATATCTTGGAGAAATCATGGNAGATATGAGATATT 2811  
2698 CTGAGCAGCAAGCAGCGCGCGCTCTGTGAGGCGCCACCGAGAGAGGACTACGGGTGCTG 2756  
2812 AGACAGGACCAATATGATGCTCTCAAGGAGGAGATACCTTACCCTTTAAGTCAAGCCAGA 2871  
2757 --GCGGAGACGGCACCGCCAGAGAAAGAGCGCGCGGATGTGAGCTGGAAAGGCCCGC 2814  
2872 AAAAGTGGTTTCCAAATGGAATGGCTGTCTGAACCTTCAAGGAGGAGGAGGAGGTTATT 2931  
2815 GCCAACCGGACGCGATCGACTGGGACGCGCTACACGCGCGGTGCGCGCGATCGGTGCG 2874  
2932 GGGACCCAGGTCTTTGAAGACTATGACCTGAGAGAGCTGGTGAGCTACATTTGACTGGAAG 2991

2875 GGATACGGGAATTTAGGACTACGACTCGCGAGCTCGCGAGTACATCGACTGGCAG 2934  
2992 CTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCCAAGATA 3051  
2935 CCGTCTTCAACGCTGGGAGAT-----GAAGGTCTGCTTCCCCGACATC 2979  
3052 TTAAACGAAACACAGTAGGTGGAGAGGCCAGAGGCTCTACGATGATGCCCAATATG 3111  
2980 CTCAACAAACCGGCGACCGGCGAGGCGGCCCGAAGCTCTACGACGACGCCAGGAGATG 3039  
3112 CTGAACACACTGATTAAGTCAAAAGAAATCTCGGCGCCCGGGTGTGGTGTCTTGCCCA 3171  
3040 CTGACACACCTGATCAAGGAGAGTGGCTGACGGCCAAACGCGGTGATCGGTTCTTCCCG 3099  
3172 GCACAGAGTATCAAGACGACATTCACCTGTGA GCGAGAGCTGTGTGCCCCAGAGGTGCA 3231  
3100 GCGAACGCGATCGCGCGCGGTTTCGAAGACATCGAGCTGTACACCGACGACACCGCAC 3159  
3232 GAGCCCATAGCCACTTCTTATGGGTTAAGGCAACAGCTGAGAGGATCTTCCGACGACG 3291  
3160 GAGGTGCTGACCAACGCTGCAACCTGCGCGCAGAGGGCGAGCACCGCGACGGAATC--- 3216  
3292 GAGCCATATCTATGCTCTCAGACTTTCATCGCTCCCTTGTGATTTCTGGCATCCGTGACTAC 3351  
3217 ---CCGAAACCGGTGCTGGCGGACTAGTGTGCGCCCAAGAAACGGGCAACCGGACTAC 3273  
3352 CTGGGCTGTGTTCCGCT---TGCTGCTTTGGGTAGAAGAGCTGAGAGCTGAGAGCTATGAG 3408  
3274 ATCGGCGCTTTCGCGGTGACCGCGGGCTGGCGAGCGAGAGATTCGGGATTCNAG 3333  
3409 GATGATGTTGACCACTACAGCAGCATATGTTCAAGCGCTGGGGGACCGGCTGGCAGAG 3468  
3334 CGGCGCTTGGACGACTACAGCGCGATCTGCTGAGGTGCGATCGCCGACCGGCTGGCGAG 3393  
3469 GCCTTTGCAAGAGCTCCATGAAGAGTTTCGCGGAGAACTGTGGGCTTACTGTGGCAGT 3528  
3394 CGGTTGCGCGAAGAGTGCACAGCGGGTCCGCAAGAGTTCTGCGGCTACAGCCCGAC 3453  
3529 GAGCAGCTGAGCTCGCAGACCTTCGGAAGGTTTCGGGTACAAAGGGCATTCGCCCGGCTCCT 3588  
3454 GAGCAACTGCAACAAACGCGCATCATCGACGAGAGTACCGGGGAATTCGCGCGCGCG 3513  
3589 GGCTACCCAGCAGCGCGGACCAACGAGAACTCACTGTGAGAGCTCGCGAGACATC 3648  
3514 GGCTACCGCGCTTCCCGGAGCACCAGAGAGGTGACGCTGTGGAAGTTGATGAGAGTC 3573  
3649 GAGCAGTCTACAGCATTAGTTAAACAGATCATTAGCATGCACTGCTTTCAGCAGTC 3708  
3574 AAGAGCGCACCGCATCGAGCTGACGAGTGCATGCGCCATGTGGCCCGCGCGCGCTC 3633  
3709 TCAGGCGCTTACTTCTTCCAAATTTGAAGTCCAAATATTTTCTGTGGGAGAGATTTCCNAG 3768  
3634 AGCGTTGGTATTTCTCGCACCGCGAGTCGAGTCTTCTGTGTGCGGCGGTGGCCAG 3693  
3769 GATCAGTTGAGAGATTTATGATTTGAGAGAGAAATATCTGTGCTGTGAGTTGAGAAATGG 3828  
3694 GACCAGGTGCGGACTACGCGAAGCGCAGAGGCTGAGCGCTGCGCGAGGCGCGAGCTG 3753  
3829 CTTGGACCACTTTTGGGATATGATACAGA 3857  
3754 CTGGCCCCCAACCTCGGCTACAAACCCCGA 3782

## RESULT 4

US-10-282-122A-8842  
; Sequence 8842, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari

; APPLICANT: Zykkind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Foreyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8842  
; LENGTH: 3684  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-10-282-122A-8842

Query Match 23.3%; Score 911.8; DB 7; Length 3684;  
Best Local Similarity 55.3%; Pred. No. 3.3e-264;  
Matches 2070; Conservative 0; Mismatches 1557; Indels 114; Gaps 11;

Qy	128	TCATGCCATCTTCGACAGAGGATTATGCTGCTGGATGGAGGATGGGACCATGATCC	187
Db	20	TAAAGCCCTTCTTGCTAAACGCATCTCGATTATGATGGTGAATGGGAACCATGATCC	79
Qy	188	AGCGGGAAGCTAAACGAAGAACATCTCCGAGGTGAGGATTTAAAGATCATGCCAGGC	247
Db	80	AGCGCCATAATTTGGAAGAGCTGACTATCGTGGTGAAGCTTTGCTGATTTGGGCACATG	139
Qy	248	CGCTGAAAGGCAACATGACATTTTAAAGTAACTACCTGACCTGATGCTATTTACCAATCC	307
Db	140	ACTTAAAGGTAACAAATGACCTTTTGGTCTTAAACACAGCCTCAATCATCTCAAGGTATTC	199
Qy	308	ATAGGATATCTTGCTGGCGGAGATATCATTTGAAACAAATCTTTTACGAGCACTA	367
Db	200	ATGAAGCTCACTCGATGCTGGTGAGACATTTATTTGAAACCAACAGCTTTTAAACGGTACAC	259
Qy	368	GTATTGCCAAGCTGACTATGCTTTGAACACTTGGCCCTACCGATGAACATGTGCTCTG	427
Db	260	GTGTTTCAATGCTGACTATCACATGGAAGATCTTGTTCCAGAGATTAAACCTGAGCAG	319
Qy	428	CAGGAGTGGCAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGAAATTAAGAGTTTG	487
Db	320	CACGTTTATAGCTTAAAGCAGCTTGGCAAG--TATTCAACGCCACACAAAGCTCGTTTG	376
Qy	488	TGGCAGGGCTCTGGGTCGACTATAAGACACTCTCTGTCTCCCACTCTGTGGAAAGGC	547
Db	377	TGGCAGGGTGTACTTGGGCCAATCTCGTACATGTTCAATCTCTCCAGATGGAACACC	436
Qy	548	CGGATTATAGGAACATCATCTTGTATGAGCTTGTGTTGAAAGCATACCAAGAGCGAGGCAAG	607

Db	437	CTGCTTTTCGTAACATTTAGCTTTTGTGTAACATAAAGAGAAATATATTTGAAGCGACTCATG	496
Qy	608	GACTTCTGGATGGGGGTTGATATCTTACTCATTTGAAACTATTTTGTGACTGCCAATG	667
Db	497	CACATATTGAAGTGGTGCAGACATATCTCTGATCGAAACTGATTTGATTTAAACT	556
Qy	668	CCAAGGCAGCTTGTTTGCACTCCAAATCTTTTGTAGGAGAAATATGCTCCCGGCTA	727
Db	557	GTAAAGCAGCGATTTTGTGAGTCAGTCAAGAGTCTTTAAACAAATTTGGTCGGAATTACAA	616
Qy	728	TCCTTATTTTCAGGAGCGATCGTTGATAAAAGTGGGGGACTCTTTTCGGGACAGACGAG	787
Db	617	TTATGATTTTCAGGACCATTTACCGATGTCATCAGGCCGTACTTTAAACAGGTTCAGACAGCG	676
Qy	788	AGGATTTTTCATCAGCGTCTCTCATGGAGAACCACTCTGCATTTGATTTAAATTTGCTT	847
Db	677	AAGCTTTCTGGAACCTCGGTTTCGTCATGGCGCATTTGCTTTCAATCGGTTTAACTGTGCC	736
Qy	848	TGGTGCAGCTGAGATGAGACCTTTTATGAAATATTTGAAATATGTAACACAGCCTATG	907
Db	737	TTGGTGCAGATGCCATGGCCCTCACGTAAACAACTATTTCCGATGTCGAGATACCTTTG	796
Qy	908	TCCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTTGGTGTGACTATGATGAAACGCTT	967
Db	797	TTTCAGCGCACCCAAATGCAAGGCTTACCAACGCAATTTGGTGAATATGACGAAACTCCAG	856
Qy	968	CTATGATGGCCAGCACCTTAAAGATTTTGTCTATGGATGCTTGGTCAATATAGTTGAG	1027
Db	857	AGCAAACTGCAGCTTTCTTAAAGAGGTTTCTGTAAGCGGTTTGAATTAACATTACTCGTG	916
Qy	1028	GATGCTGTGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAATTTGTA	1087
Db	917	GTGCTGTGTACGACACCAAGACCATATTCGAGCTATTCGCAATTCGCTTAAGACATTTG	976
Qy	1088	AGCTAGAGTTTCCACCTGCTTTTGAAGGACATATGTTACTCTCTGCTGTAGAGC	1147
Db	977	CGCTCGCAAGTGCCTGAAACCGTACCTGCTTGC-----GCTTAAGTGGTTAGAAC	1030
Qy	1148	CTTTCAGGATGGACCGTACCAACCTTTGTTAACTTGGAGGAGCGCTGTAATTTTCAG	1207
Db	1031	CATTTAATATTTATGATGATTTCTTAAACGTTGGGAGCGTACTAACCTTTACC	1090
Qy	1208	GATCAAGGAAAGTTTGTAAACTCATCATGTCAGGAACTATGAAGAGCCTTTGTGTGTG	1267
Db	1091	GTCTTAAATTTCTTACGCTCTCATTCGTGAAGAAACTTTGCGAAGCTTTAGAAAGTTG	1150
Qy	1268	CCAAAGTCAGGTGGAAATGGAGCCAGGTGTTGGATGTCAACATGGATGATGGCATGC	1327
Db	1151	CACAGCAGCAGGTGGAAGCTGGCGCACAGATTTATCGACATTTAAACATGGATGAAGGATGC	1210
Qy	1328	TAGATGTCGATGCAATGACAGATTTTGCACCTTTAATTTGCTTCCGAGCCAGACATCG	1387
Db	1211	TCGATTCGCAAAATGCGATGCTGCTTTCTTAAACCTTTAGTACATTCGAAACCGGACATTT	1270
Qy	1388	CAAGGTACCTTTGTGATCGACTCTCCAAATTTTGTCTGATTTGAAGAGCTGGGTTAAAGT	1447
Db	1271	CACGTGATCCGATCATGATTTGACTCATCGAAATGGGAAATCATTTGAAGCCGGCTTAAAT	1330
Qy	1448	GCTGCCAAGGAAAGTGTCAATAGCATTTAGTCTGAAGGAAAGGAGAGAGCACTTCT	1507
Db	1331	CGGTACAAAGGTAAACCGGTTGTTAACTCAATTTCTTAAAGAAAGGTTATGACGAGTTTG	1390
Qy	1508	TGGAGAGGCCAGGAAGATTTAAAGATTTGAGCTGCTATGGTGTGCTATGCTTGTGATG	1567
Db	1391	TTGAAAGGCCCGCTCTGCGTCAATATGCTGCTGCAATTTATTTGTGATGCTTTTACG	1450
Qy	1568	AAGAGGACAGGCAACAGAACACACAAATAACAGAGTGTGCACCCCGGCTTACCATC	1627
Db	1451	AAGTATGTCAGGCCGACACTGCTGAAACGTAACCTGTAACCTGTAGCGCTCTTATGACA	1510
Qy	1628	TGCTTGTGAAAAAACTCGGCTTTTAAATCCAAATGACATTTATTTTGTGACCTTAATTCCTAA	1687

Db 1511 TTTTGGTTAAACGAAGTAGGCTTCCCTGCTGAAGATATTATTTTGGACCGGAACGTGTTG 1570  
Qy 1688 CCATTGGGATGGAGGAAACAACTCTGTATGCCATTAATTTTATCATGCAACAA 1747  
Db 1571 CAGTTGGGACTGGTATTGAAGAAACATACAACTACGCGATCGATTTTATTGAAGCAACGG 1630  
Qy 1748 AAGTCATTAAAGAACATTACCTGGAGCCGAATAAGTGGAGGCTTTTCCAACTGTGCT 1807  
Db 1631 GCTGGATTAAACAGAACTTACCGCAGCCATGATTTCTGGGGTGTGCTAACGTTTCGT 1690  
Qy 1808 TCTCTCTCCGAGGAATGAAGCCATTCGAGAAGCAATGCAATGGGGTTTTCCTTTACCATG 1867  
Db 1691 TCTCATTCGCTGGCAATGAACCACTTCTGAGGCCATCTACTCTGTATCTTGTACCATG 1750  
Qy 1868 CAATCAAGTCTGGCATGAGCATATGAGATAGTGAATCTGGAACCTCCCTGTGTATGATG 1927  
Db 1751 CCATCAAGCAAGGCATGACCATGGGTATTGGAAGCGAGTCAATGGCTATTATGATG 1810  
Qy 1928 ATATCCATAGGAATCTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGACCTGAGG 1987  
Db 1811 ATATTCTTACCGGCTAAAGAAAGCGGTGAAGATGTCATTTTAAACCAAAATCAAGGTG 1870  
Qy 1988 CCATCGAAGACTCTTACGTTATGCCCCAGACTCAAGGCACA----- 2028  
Db 1871 AGTCTGTCAAGTCCGACTGAGAACTACTTGAAGTTGCGAATAATACCGTGGACAAG 1930  
Qy 2029 --GGAGGGAAGAAAGTCAATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGACGCC 2086  
Db 1931 GTGGTGCAACAAAGAACCGAGAACCTTGAATGGCGTAAATGAGTCAGTGTGAATAACGTC 1990  
Qy 2087 TTGAGTATGCCCTTGTGAGGGCAATGGAAGCAATATTTATGAGGATCTGAGGAAGCCA 2146  
Db 1991 TTGAATATGCCCTTGTGTTAAAGGTATTACGACTTATATTGACCAAGACACCGGAAGGCC 2050  
Qy 2147 GGTAAACCAAAATAATATCCCGACTCTCAATATAATTTGAAGGACCCCTGATGAATG 2206  
Db 2051 GCTTAAATCAAAA-----CGTCTTTAGATTAATTTGAAGGCGCACTGATGACG 2101  
Qy 2207 GAATGAAATTTGTTGGTATCTTTTGGAGCTGGAATAATTTTCTACCTCAGGTTATAA 2266  
Db 2102 GCATGAAGCTGCTGGTGACTTGTTCGGTTGAGGCAAAATGTTCTTGCCACAAGTTGTA 2161  
Qy 2267 AGTCAGCCGGTATTGAAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAA 2326  
Db 2162 AATCTGCCGAGTCAATGAACAAAGCAGTAGCATGGCTCAACCCGCTACATCAGAGCTGAA 2221  
Qy 2327 GAGAGAAACCAAGATGCTTAAACGACAGTAGAAGAGAGGACCTTACAGGGCACCA 2386  
Db 2222 -----AGACCGAAGGACAGTCTTAAAGGTAAAG 2248  
Qy 2387 TCGTGTGGCCACTGTTAAAGCGACGTGACGACATAGGCAAGACATAGTGGAGTAG 2446  
Db 2249 TCCTAATGGCAACGGTTAAAGTGACGTACACGATATTGGTAAATAATTTAGCGGTAG 2308  
Qy 2447 TCTTGGCTGCAATAATTTCCGAGTTATGATTAGGAGTCAATGACTCCATGTGATAAGA 2506  
Db 2309 TACTTGGCTGTAATGGCTATGACATTTGTGACCTTGGCGTAATGTCCTTTCGAGAAA 2368  
Qy 2507 TACTGAAAGCTGCTTGAACACAAAGCAGATATAATTTGGCTCTGAGGACTCATCTC 2566  
Db 2369 TCTTACAACTGCAATTGATGAAAAATGTGACATCATTTGAGTCTGTGCTGATCACCC 2428  
Qy 2567 CTTCCTCGATGAATGATTTTGTGCAAGGAAATGAGAGATTAGCTATAGGATTC 2626  
Db 2429 CATCTTTAGATGAATGATTTTGTGCTAAGAAATGACGGTAAAGGCTTTTAAACATTC 2488  
Qy 2627 CATTTGTTGATTTGGAGGACCAACCTTCAAAAAACCAACAGCAGTAAAAATAG---CTC 2683  
Db 2489 CTTTATTAAATTTGGTGGCGGACTACTTCTAAGACACATACAGCAGTAAAAATTTGACCTC 2548  
Qy 2684 CGAGATACAGTGCACTGATATCCATGCTGAGCGGTCCAGAGTGTGGTGTGTT 2743  
Db 2549 AGTATCAAAACGATGCGGTAAATTTATGTTGCCGATGCTTACGCTGCTTGTGCTAGCGA 2608

Qy 2744 CCAGAGCTGTAGATGAATCTTAAAGATGAATCTTTGAGGAATCATCGAAGATATG 2803  
Db 2609 CAACCTTGTCTTCCGAAGAAATCGTGGAGCATTTATTGAAGAGCATCTGCTGCTGAATATG 2668  
Qy 2804 AAGATATTAGACAGGACCATTTATGAGTCTCTCAAGGAGAGGAGATCTTACCTCTTAAGTC 2863  
Db 2669 CCAAAATCTGTGAGCTTTAGCCAAACAACAACCAAAAGCAGCCAACTGACTTTATAAG 2728  
Qy 2864 AAGCCAGAAAAAGTGTTCCTCAAAATGATGGTGTCTGAACCTCACCCAGTGAAGCCCA 2923  
Db 2729 AGTCGGTTGAATAATGGTTTAAATAATGATGAAGACTACGTGCCAC-----CAAAAACAA 2782  
Qy 2924 CGTTTATTTGGACCCAGGCTTTTGAAGACTATCACTGACAGAGCTGCTGAGCTACATTTG 2983  
Db 2783 ATCTTTTGGGAACACAAAGTTTAAAGAAATTCGCTTGTCTACACTGCTGGAATTTATTTG 2842  
Qy 2984 ACTGGAAGCCTTTCTTTGATGCTGTGGCAGCTCCGGGGCAAGTACCAGAACTGAGGCTTCC 3043  
Db 2843 ACTGGAAGCCTTTCTTTTATTTCTTGGAGTTTAACTGGCAA-----ATTCC 2887  
Qy 3044 CCAAGATATTAAACGACAAACAGTAGGTGAGAGGCGCAGGAAGGTCTACGATGATGCC 3103  
Db 2888 CGAAAATTTTGAAGATGAAGTGTGGCGAAGCAGCAACTGACCTTTTACAACCCAGGCAC 2947  
Qy 3104 ACAATATGCTGAACACACTGATTTAGTCAAAAGAAACTCCGGGCGCGGGGTGTGGTGGGT 3163  
Db 2948 AAGCGATGTTGAAGATATTATTCGACAAACAATCGTTTGTATGCTGCTGCTGATTTGGTA 3007  
Qy 3164 TCTGGCCAGCACAGAGTATCCAAAGACGACATTCACCTGTACGAGAGGCTGTGTGCCCC 3223  
Db 3008 TGTTCCCTGCTCAGCGGTACAGATGCGAGATACCGTACGCGTATTTGATGAAGCTGTCAAA 3067  
Qy 3224 AGGCTGACAGACCCATAGCCACATTTCTATGGGTTAAGGCAACAGGCTGAGAAGACTCTG 3283  
Db 3068 ATGTTAGCCA-----TACTTTGAGCAGCTTTACCCAGCAATCTGACAGGTGACAG 3118  
Qy 3284 CGAGCAGGAGCCATACTACTGCTCTCAGACTTCACTTCATCGCTCCCTTGCATCTTGGCATCC 3343  
Db 3119 GCA-----AACCAAAATTTATCTTGGCAGATTACAT---TCGAGCTGATCGGAGCAGC 3169  
Qy 3344 GTGACTACTTGGGCGCTTTTGGCGTTGCTGCTTGGGTTAGAGAGCTGAGCAAGGCT 3403  
Db 3170 AAGACTACTTTGGCGGATTCACCTGTATCGATTTTGGTGAGAAAGAACTGGCAATGAAT 3229  
Qy 3404 ATGAGGATGATGTGACGACTACAGCAGCATCATGTTCAAGGCGCTCGGGAGCCGGCTGG 3463  
Db 3230 ACAAAGCCAAAGTGATGACTACTCTGCAATTTTGTGAGTCAATTAGCTGACCGTTTG 3289  
Qy 3464 CAGAGGCTTTTGAGAGAGCTCCATGAAAGAGTTCCCGAGAACTGTGGGCTTACTGTG 3523  
Db 3290 CCGAAGCTTTTGGCGAACACTTACATGAAGCTATTCGTAAGAGTTCTGGGGCTATAAAG 3349  
Qy 3524 GCAGTGAAGAGCTGAGAGCTGCGAGACCTGCGAAGGTTGCGGTACAGGGCATCCGCCGG 3583  
Db 3350 CTGATGAGCAGCTCAGCAATGAAGAACTGATTAAGAGAAATATGTCGTTATTCGCCCTG 3409  
Qy 3584 CTCTGCTACCCAGCCAGCCAGCACACCCAGAGAGCTCACCATCTGGAGACTCCGAG 3643  
Db 3410 CACAGGCTACCTGCTTGGCCAGAGCAGCTCTGAAAAGCAGGTGTTGTTGACTGGTTAG 3469  
Qy 3644 ACATCGAGCAGCTTACAGGCAATTAGGTTAAACAGAACTATTAGCAATGGCACTGCTTCAG 3703  
Db 3470 GTTCTACCGACAAAATCGGCACCAACTGACTGAGCACTTTCGATGATCGCGCACTTT 3529  
Qy 3704 CAGTCTCAGGCTCTACTCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTT 3763  
Db 3530 CAGTAAGTGGTTCTATTATCTCATCTCAGAGTGAATACTTTAACTGGGTAAAAATCT 3589  
Qy 3764 CCAAGATCAGGTTGAGGATTTATGCAATTGAGGAGAACATATCTGTGCTGAGGTTGAGA 3823  
Db 3590 CTCAAGACCAACTTGAAGATTTATGCAAAACGTAAGGTTGGACACTGGATGAAGCGAAGC 3649









Db 3427 ACAGAAAAGCAAAATTTGGCAACTGTTAAATGTAGAAAATCGCATTTGGTATGAAGCTT 3486  
Qy 3673 ACAGAAATCATTTAGCAATGGCACTGCTTCAGCAGTCTCAGGCCCTCTACTCTCCAAATTTG 3732  
Db 3487 ACTGATGCTTACGCGATGTGGCCGGGGCTTCGGTATCGGGTTGGTATTTTAGTCATCCT 3546  
Qy 3733 AAGTCCAAATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTTGAGGATTTATGCATTG 3792  
Db 3547 GAAAGTAATATTTTGGCGTGCCTCAAAATTCAAAAGATCAGGTAGAGATTAAGCCAAA 3606  
Qy 3793 AGGAAGAACATATCTGTGGCTGAGGTTTGAGAAATGGCTTGGACCCATTTTGGGATATGA 3851  
Db 3607 CGTCGTGTATGAGTGTGAGTGAAGTTGAGCGTTGGTTAGCCCTTAATTTAGGTATGA 3665

## RESULT 6

US-10-282-122A-30149  
; Sequence 30149, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 30149

; LENGTH: 3705

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-10-282-122A-30149

Query Match 23.0%; Score 901.4; DB 7; Length 3705;  
Best Local Similarity 55.2%; Pred. No. 4.7e-261;  
Matches 2073; Conservative 0; Mismatches 1581; Indels 103; Gaps 12;  
Qy 113 CCCTCGGGATGAGATCAATGCCATTTGCGAAGAGGATTTAGTGTGATGAGGGA 172  
Db 29 CCGCGCTGCAAGCCCTCCAGCAGCGCCTCAGGGAACGTATCCTGATCCTCGATGCGGCA 88

Qy 173 TGGGACCATGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTGAGGAATTTA 232  
Db 89 TGGGACCATGATCCAGAGCTAAGCTGGAAGGCGGACTACCGGCGGAGCGCTTCG 148  
Qy 233 AAGATCATGCGAGCGCGCTGAAAGGCAACATGACATTTTAAAGTATTAACCTCAGCTCATG 292  
Db 149 CCGACTGCGCGAGCGAGCTGAAAGGCAACAAACGACCTCTTGTGCTGAGCGCGCGGACG 208  
Qy 293 TCATTACCAAAATCCATAAGGAATACTTGTGCTGGTGGGCGAGATATCATTTGAACAATA 352  
Db 209 TGATCCAGGGCATCGAGAAGCCCTTACCTCGACGCGCGCGGACATCTCAGAGCAACA 268  
Qy 353 CTTTTCAGCAGCTAGTATTTGCCCAAGCTGACTATGSCCTTGAACACTTGCCTTACCGGA 412  
Db 269 CTTTCAAGCCACCCAGGTGTCCAGGCCGACTACGGCATGCGAGTCTGCTGCTTACGAAC 328  
Qy 413 TGAACATGTCTCTGCAAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGA --- 469  
Db 329 TCAACGTGGAAGGGCGCGCTGCGCCGCCAGGTGGCGGAGCGGAAGACCGCGAGACCC 388  
Qy 470 CAGGAATTAAGAGGTTTGTGCGAGGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGT 529  
Db 389 CGGACAAAGCGCGTTTGTGCGCGGCGTGTCTCGGCCCGACCGACCGCCTGCTCGATTT 448  
Qy 530 CCCCATCTGTGGAAGGCGCGGATTTATAGGAACATCAACATTTGATGAGCTTGTGGAAGCAT 589  
Db 449 CCGGAGCGTGAACACACCCCGGTACCGCAACGTCACTTCGACGAACTGGTGGAGAACT 508  
Qy 590 ACCAAGAGCGGCCAAGAGACTTCTGGATGCGCGGGTTGATATCTTACTCATTTGAAACTA 649  
Db 509 ACGTCGAGGCGACCCGAGGCTGTATCGAAGCGCGCGCGACCTGATCCTGATCGAGACCA 568  
Qy 650 TTTTGTGATCTGCAATGCGAAGGAGCGCTTGTGTCACCTCCAAAATCTTTTGGAGAGA 709  
Db 569 TCTTCGACACCTCAACGCCAAGCGCGCGATCTTTCGCGCTCCAGGGCGTGTTCAGGAAC 628  
Qy 710 AATATGCTCCCGCGCTATCTTTATTTTTCAGGAGCATGCTTGTGATAAAGTGGCGGACTC 769  
Db 629 TCGGCGTGGAGCTGCCGATCATGATCTCCGAAACATCACGAGCGCTCCGCGCGACCC 688  
Qy 770 TTTCCGACACAGACGAGAGAGGATTTGTCTATCAGCGTGTCTCATGGAGAACCACTCTGCA 829  
Db 689 TGTGCGGCCAGACACCGAGGCTTCTTGGAACTCGGTGCGGCATGCCCGCGATCTCGG 748  
Qy 830 TTGATTTAAATTTGCTTTTGGGTGACCTGAGATGAGACCTTTTATTTGAAATAATTTGAA 889  
Db 749 TAGGCTTGAATGCGCGCTCGCGCCCAAGGAAATTCGCGCGGTATCATCGAGGAACCTGCGA 808  
Qy 890 AATGTACACAGCCTATGCTCTGTTATCCCAATGCGAGGTCTTCCCAACACCTTTGGTG 949  
Db 809 CCAAGGCGGACACTCATGCTCGGCCCAACCCCAAGCGCGGCTGCCGAACGCTTCGCG 868  
Qy 950 ACTATGATGAAACCGCTTCTATGATGGCCAAAGCACCTTAAAGGATTTTGTCTATGATGGCT 1009  
Db 869 AATAGCAGGATCGCGCGGGAATGCGCGTGTGTTGAGGATTCGCGCGCGCGCT 928  
Qy 1010 TGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTG 1069  
Db 929 TCCTCAATATCGTGGCGGCTGCTGCGGCACACCCCGGCGCATATCGAGCGATCGCCA 988  
Qy 1070 AAGCTGTGAAAATTTGAAGCTTAGAGTTCACCTGCCACTGCTTTTGAAGGACATATGT 1129  
Db 989 AGGAGTGGCCCAAGTACCCCGCGCGGCGCATTCGCGAGATTCGCCGCGCTGTC----- 1042  
Qy 1130 TACTGTCTGGTCTAGAGCCCTTACAGATTGAGCCGTCACACCAACTTTTGTAACTTTGAG 1189  
Db 1043 GCCTGTGGGCTGAGCGGTTTACCATCGACCGAGCTCGCTGTTCTGCAACGTCGCG 1102  
Qy 1190 AGCGCTGTAATGTTGCGAGGATCAAGGAGGTTTGTGTTAACTCATCATGCGAGGAACCTATG 1249  
Db 1103 AGCGACCAACATCACCGGTTTCGCGCAAGTTTCGCGCGCGTGTATCCGCGAGGAAACTACG 1162

QY 1250 AAGAGCCTTGTGTGTTGCCAAAGTGCAGGTGGGAAATGGGAGCCAGGTGTGGATGCA 1309  
Db 1163 CGGAAGCTCTCGAGGTGGCGCTCAGCAGAGTGGGAAGCCGGCGCCAGGTATGACATCA 1222  
QY 1310 ACATGATGATGGCATGCTAGATGGTCAAGTGCAGTCAAGTGCAGATGATGCAATTTTGCACATTAATTG 1369  
Db 1223 ACATGGACGAAGGATGCTGACTCGAAGGCGGCATGCTCACCTTCTCAACCTGATCG 1282  
QY 1370 CTTCCGAGCCAGACATATGCAAAAGTACTTTGTGCGATGCGATCTCTCCAAATTTTGTCTGTGA 1429  
Db 1283 CTTCCGAGCCAGACATCTCGCGCTGCGCATCATGATCGACTCTCTCCAAAGTGGGAAGTGA 1342  
QY 1430 TTGAAGCTGGGTTAAAGTGTGCCAAGGAGTGCATTTGCAATAGCATTTAGTCTGGAAGG 1489  
Db 1343 TCGAGGCGGGCTGAAGTGCATCCAGGCAAGGCGATCGTCAATCGCATCTCGAATGAAGG 1402  
QY 1490 AAGGAGAGGAGCACTTCTTGGAGAGGCGCAGGAAGATTTAAAGATATGAGAGCTCTATGG 1549  
Db 1403 AAGGGTTCGAGGCTTCAAGCACCATGCGCGCTGTGCAGGCGCTACGGCGCGCGGTGG 1462  
QY 1550 TGGTCATGGCTTTTGTATGAAGAGGACAGGCAACAGAAACAGACACAAATATCAGATGT 1609  
Db 1463 TGGTATGGCTTCGACGAGGAGCGCCAGGCGCGACACCCAGGCGCGCAAGGAAGAAATCT 1522  
QY 1610 GCACCCGGGCTACCATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATTTT 1669  
Db 1523 GCAAGCGCTCTACGACATCTGTCGACGAAGTCCGCTTCCACCCGGAAGACATCATCT 1582  
QY 1670 TTGAACCTTAATCTTAACATTCGGACTGGAATGGAGGACACAACTTGTATGCCATTA 1729  
Db 1583 TCGATGCGAATCTTTCGCAATGCCACCGCATCGAGGAACACAACTACGCGGTG 1642  
QY 1730 ATTTTATCCATGCAACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAG 1789  
Db 1643 ATTTTATCAACGCTGGCGCTTACATCGCGACAACTCCCTACGCTCGAGCTCGGGCG 1702  
QY 1790 GTCCTTTCGAATCTGCTTCTTCTTCCGAGGAATGGAAGCATTCGAGAGCAATGATG 1849  
Db 1703 GGGTGTCCAACTGCTCTTCTGTTCCGCGCAACAAACCCGGTACGCGAGCGATCCACT 1762  
QY 1850 GGGTTCCTTTACATGCAATCAAGTCTGCAATGGAATAGTGAATGCTGGA 1909  
Db 1763 CGGTGTTCTCTACTACGCGATCGCAACGGCTTGACATGGGCATGTCACGCCCGCC 1822  
QY 1910 ACCTCCCTGTGTATGATGATATCCATAGGAATCTTCTGACGCTCTGTGAAGATCTCATCT 1969  
Db 1823 AGCTGGAATCTACGACGAGATTCCGAAAGCGCTCGCGACCGGGTCGAGGACGTGTGC 1882  
QY 1970 GGAATAAGACCTTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCA-----GGCA 2026  
Db 1883 TCAACCGCACCGCCGAGGCCACCGAGGCCCTGCTGGCGATCGCGATCGACGACTACAAGGGCG 1942  
QY 2027 CAGGAGGGAAGAAAGTCAATTCAGACTGATGATGGAAGAAATGGCCCTGTGGAAGAGCC 2086  
Db 1943 CGGCGCGGTCAAGAGGCGGAGGACGAGGAATGGCGAGCTACAGGCTCGAAGAGCGCC 2002  
QY 2087 TTGAGTATGCCCTTGTGAAGGCGATGGAAGAAATATTTATGAGGATCTGAGGAAGCCA 2146  
Db 2003 TCGAGCATGCGCTGTCNAAGGCGATCACCCTCGATCGTCTGAGACACCGGAG 2057  
QY 2147 GGTTAACCAAAAAATATCCCCGACTCTCAATAATAATTTGAAGGACCCCTGATGAATG 2206  
Db 2058 -----ATGCCGCGCAGCAGTGTGCGGTCCCATCGAGGTGATCGAAGGTTCGCTGATGTCG 2113  
QY 2207 GAATGAATAATTTGTTGATCTTTTTCGAGCTGGAAGAAATTTTCTACTCTCAGGTATTA 2266  
Db 2114 GGATGAACGTGTTCGCGCGACTGTTTCGCGCGCGGCAAGATTTCTCTCCGAGGTGTCA 2173  
QY 2267 AGTCAGCCCGGTTATGAAGAGGCTGTGCGCCACCTTATCCCTTTTCATGAAAAAGAA 2326  
Db 2174 AGTCGCGCGAGTGTATGAGCAGGCGGTGGCCCACTGATTCCTTTCATCGAGGCGGAGA 2333  
QY 2327 GAGAAGAAACAGAGTGTCTTAACGGCACAGTAGAAGAGAGGACCTTTACCGAGGGACCA 2386

Db 2234 AAGCGACA-----AGCGGAAGCCAGGGCAGA 2263  
QY 2387 TCGTGTGGCCACTGTTAAAGGCGAGTGTGCACGACATAGCAAGAAACATAGTTGGAGTAG 2446  
Db 2264 TCGTGTGGCCACGGTGAAGGCGGAGTGTGCACGACATCGGAAGAAACATCGTGGCGTGG 2323  
QY 2447 TCGTGTGGCTCAATAATTTCCGAGTTATTGATTTAGGAGTCATGACTCCATGTATGAAGA 2506  
Db 2324 TCGTGTGGCTCAACCGGTATGACGTGGTTCGACCTCGGCGGTGATGGTGGCGGCGAGAAGA 2383  
QY 2507 TACTGAAAGCTGCTCTTGCACCAAAAGCAGATATAATTTGGCCTGTGAGACTCATCACTC 2566  
Db 2384 TCGTGCAGACCGCATCGCGAGAAATGCGACATCATCGSCCTGTCTGGCTCTGATCACGC 2443  
QY 2567 CTTCCCTGGATGAATAATTTTGTGCAAGAAATGGAAGATTAGCTATTAAGGATTC 2626  
Db 2444 CGTGTGTGGACGAGATGGTCCAGTGCACAAAGAAATGCAAGCGCAGAAATTTCCAGTTGC 2503  
QY 2627 CATTTGTTGATGGAGGAGCAACCACTTCAAAAACCCACACAGCAGTGTAAATAGTCCGA 2686  
Db 2504 CGCTGATGATCGGCGCGCCACTACCTCGAAGCGCATACCGCGGTGAAGATCGATCCGC 2563  
QY 2687 GATACAG-----TGCACTGTAAATCAATGCTCTGACGCGTCCAAAGAGTGTGTGTGTT 2743  
Db 2564 AGTACAGCAACGACGCGTGGTCTACGTCAACGCGCTCGCGCGGTAGGCGTGGCCA 2623  
QY 2744 CCAGCTGTTTGTAGATGAATAATCTAAAGGATGAATTAATTTGAGGAATCATGGAAGATATG 2803  
Db 2624 CCAGCTGCTGTCCAGGAGTGAAGGCGGACTACGTGGCCCGCACCCCGCGGACTACG 2683  
QY 2804 AGATAT-----TAGACAGACCATTAATGATCTCTCAAGGAGAGAGATACTTACCCTT 2858  
Db 2684 CGGTGTGTCGCGAAGCAGCAGCGCAACCGCAGCGCGCCACCGAGCGGCTGAGCTACGAAC 2743  
QY 2859 AAGTCAAGCCAGAAAAGTGGTTTCCAAATGGAATGCTCTGTGAACCTCAACCCAGTGAA 2918  
Db 2744 AGCGATTCGCAACAAGCCGCGGTTCGACTGGGCGGCTACCAAGCGCCGAC----- 2795  
QY 2919 GCCCACTTTATTTGGGACCCAGGCTTTTGAAGACTATGACTTCGACAGAGCTGTGTGACTA 2978  
Db 2796 GCCTTCTTCAACCGCGCTCAGGCTGCTCGACGAGATCGACTCGCGGTGCTCGCGGATG 2855  
QY 2979 CATGACTGAAGCGCTTTTGTGATGCTGCGAGCTCCGGGCAAGTACCCGAATCGAGG 3038  
Db 2856 CATGACTGAAGCGCTTTCTTCAATCTGCGACCTGGCGCGCAAGTACCG----- 2907  
QY 3039 CTTCCCCAGATATTTAAAGCAAAACAGTAGTGGAGGCGCAGGAGGTCTACGATGA 3098  
Db 2908 -----CGCATCTTCAACGACGAGTGTGGCGAGGCGCCACCTCTGTTGTTCAACGA 2960  
QY 3099 TGCCCACAATATGCTGAACACACTGATTTAGTCAAAAGAAACTCCGGGCGCGGGGTGTGGT 3158  
Db 2961 CGCCAGCGCATGCTGAGAGAGTGTGATCGACGAGAGTGTATCAAGGCGCGCGGTGTT 3020  
QY 3159 TGGTGTCTGCGCAGCAGAGTATCCAAAGCAGATTCACCTGTACGACAGGCTGTGT 3218  
Db 3021 CGGCTTCTGGCGGCCAACCCAGGTGAGACGACGACCTGAGAGTCTACGCGCGCGATGG 3080  
QY 3219 GCCCAGGCTGACAGGCCATAGCCACTTCTATGGTTAAGGCACACAGCTGAGAGGA 3278  
Db 3081 C-----GAGACCTTCGCCCTTGCACCTCGCGCAGCAGCAGATCAAGCC 3128  
QY 3279 CTCTGCAGCAGCAGGCGCATACTACTGCTCTCAGACTTCATCGCTCCCTTGATTTCTGG 3338  
Db 3129 GGACGCA-----AGCCGAACCTGCTGCTGGCGGATTCGTGCGCGCGAAGGAAGCG 3182  
QY 3339 CATTCGCTACTCTGGCGCTGTTG-----CCGTGCTGCTTTGGGTAGAGAGCTGAG 3395  
Db 3183 CGTGGCGGACTACATCGCGGCTTTCATCACCCGCGCGGATCGGCGCGGAGGAGTGGC 3242  
QY 3396 CAAGGCGCTATGAGATGATGTCAGACTACAGCAGCATCATGTTCAAGCGCTGCGGGA 3455

Db 3243 CAAGCGGTACGAAGCGGCGACGACTACACAGCATCATGTCTCAAGCGCTCGCGCA 3302  
Qy CGCGTGGCAGAGCGCTTTGAGAGAGCTCCATGAAGAGTTCGCCGAGAACTGTGGC 3515  
Db CGCGCTGCCGAAGCGCTGCGCGAGTGGCTGCAAGAGCGGTGCGAAGGAGTACTGGGG 3362  
Qy CTACTGTGGCAGTGAAGAGCTGCGAGAGCTGCGAAGGTTGCGGTACAAAGGCGAT 3575  
Db CTAGCGCGCGACGACACCTCGACAGAGCGCTTGTATCAAGGAGCAATAGTTCGGCAT 3422  
Qy CGCGCGCGCTCTGCTGCTACCCGACGCGACGACGACACCGAGAGCTACCATGTGGAG 3635  
Db CGCGCGCGACCGCGCTACCGCGCTGCGCGACGACATACCGAGAAAGCACTCTGTTCGA 3482  
Qy ACTCGGACATCGAGAGTCTACAGGCAATAGGTTACAGAACTATAGCAATAGGCAAC 3695  
Db ACTGCTGATCGGAGCGGCTGTCGGCGTCAAGCGCTGACCGAGCACTACGCGATGTTCCC 3542  
Qy TGCTTCAGAGCTCTCAGCGCTTACTTCTCCAAATTTGAAGTCCAAATATTTGCTGTGGG 3755  
Db GCGCGCGCGCTCAGCGTGGTATTTGCGCCACCGCGAGCGGCGAGTACTTCGGGTGG 3602  
Qy GAAGATTCCAGAGTCAAGTGTAGGATATGCAATGAGGAAGCAATATCTGTGGGTGA 3815  
Db CAAGATGACAAGACAGGTGGAACGCTACAGCGAGCGCAAGCGCCAGGAAGCGAGCGT 3662  
Qy GTTGGAGAAATGGCTTGACCCATTTTGGGATATGAT 3852  
Db CAGCGAGCGCTGCTGGCGCGCAACCTTGGCTACGAT 3699

## RESULT 7

US-10-282-122A-19987  
; Sequence 19987, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zykkind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19987  
; LENGTH: 3881  
; TYPE: DNA  
; ORGANISM: Enterobacter cloacae  
US-10-282-122A-19987  
  
Query Match 22.9%; Score 899; DB 7; Length 3681;  
Best Local Similarity 55.1%; Pred. No. 2.5e-260;  
Matches 2053; Conservative 0; Mismatches 1570; Indels 102; Gaps 11;  
  
Qy 149 GGAATTATGTGCTGGATGGAGGATGGGACCATGATCCAGCGGAGAACTAAACGAAG 208  
Db 44 GAATTCGTGTGGACGGCGCATGGCACCATGATCCAGGGGTATCGTCTGAGTGAAG 103  
Qy 209 AACACTTCCAGAGTCAAGAAATTTAAAGATCATGCGAGCGCGCTGAAAGGCAACATGACA 268  
Db 104 ACGATTTCCGCGCGAGCGCTTTGCCGACTGGCCCTCGACCTGAAAGGGAACACGATC 163  
Qy 269 TTTTAAGTATAAATCTCAGCGTGTATTTACCAATCCATAAGGAATATCTTGTGGCTG 328  
Db 164 TGCTGGTGTGTCAGCAAGCGCTCCGTCATTAAGGATATCCACAACGCTACTTTCGAAGCG 223  
Qy 329 GGGCAGATATCATTTGAAACAAATACATTTTAGCAGCACTAGTATTTGCCCAAGCTGACTATG 388  
Db 224 GTGGGATATGTTGAGACCAACACCTTTAACTCGACCAACCATCGCCATGCGGATTAAC 283  
Qy 389 GCCTTGAACACTTGGGCTACCGGATGAACATGTGCTCTGAGGAGTGCGCAGAAAAGCTG 448  
Db 284 AGATGGAATCCCTGTGCGCGGAATCAACCTGGAAGCGCGAAGCTGGCGCGCGCTGCG 343  
Qy 449 CCGAGGAGGTAACTCTCCAGACAGAGGAATTAAG--AGGTTTGGCAGGGGCTCTGGGTC 505  
Db 344 CCGACGAGTGGAGCGACGTACCGCGGACAAAGCGCTGCTACGTGCGGGGTGCTTGCC 403  
Qy 506 CGACTAATAAGACACTCTCTGTGTCCCATCTGTGGAAGCGCGGATATATAGAAACATCA 565  
Db 404 CGAGAACCGNAACCGGTGCTTTACCGAGCTCAACGACCGCGGTTCGTAAATATCA 463  
Qy 566 CATTTGATGAGCTTTTGAAGCATATCAAGAGCGGCAAGGACTTCTGATGGCGGGG 625  
Db 464 CTTTCGATCAGCTGTTGCGGCTTACCGTGAATCGACCAAGCGCTGTTGAAGGCGGTT 523  
Qy 626 TTGATATCTTACTGATGAACATATTTTGTATCTGCAATGCGCAAGCGACCTTGTG 685  
Db 524 CCGATCTGATCTGATAGAAACCGTATTCGACACCTTTAAGCCAAAGCGCGATTAACG 583  
Qy 686 CACTCCAAAATCTTTTGAAGGAGAAATATGCTCCCGGCGCTATCTTTATTTTCAGGACGA 745  
Db 584 CCGTGAAGAGGAGTGTGAGCGCTGGGCGTTGACCTGCGCATCATGATTTCCGCGACCA 643  
Qy 746 TCGTTGATAAAGTGGCGGACTCTTTCCGGAACAGACAGAGAGGAGATTTGTCTACGCG 805  
Db 644 TCACCGACGCTCCGCGCGTACTCTGTCCGTCAGACAAACCGAAGCGCTTTTAACTCTC 703  
Qy 806 TGTCTATGAGAACCACTCTGCAATTTGATTAATTTGCTTTGGGTGCGACTGAGATGA 865  
Db 704 TGCGCCACGCGGAAGCGCTCTCTTTGGCCCTGAACTGCGCGCTGGGCGCTGATAGCTGC 763  
Qy 866 GACCTTTTATGAAATAATTTGGAATAATGTACCAAGCGCTATGCTCTGTTTATCCCAATG 925  
Db 764 GCGAGTACGTGAGGAGCTTTCCGATCGGGATGCTAGCTACCGCCACCGGAACG 823  
Qy 926 CAGCTCTTCCCAAACACTTTTGGTGAATATGATGAACCGCTTTTATGATGCCAAGCAC 985  
Db 824 CCGGTCTGCCGAACCGCTTGGTGAATACGATCTCGATCGCACCAACCATGTCGCGCGCAAA 883  
Qy 986 TAAAGGATTTTGTATGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACAC 1045  
Db 884 TCCCGGAGTGGGCGAGTCTGGGCTTCTGAAACATCTGCGGTGGGTGCTGCGGCGACCGC 943  
Qy 1046 CAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAAAATTTGAAGCCTAGAGTTCCACCTG 1105







Db 703 TGGAACTCGGTGCTCAGCCAGCGGATTTCCGTGGGCTGAACTGGCCCTCGGCGCC 762  
Qy 856 GCTGAGATGAGACCTTTTATTTGAAATTAATTTGGAAAAATGTACAACAGCGCTATGCTCTGT 915  
Db 763 AAGGACCTGCGCCCGTATTTTGAAGAGCTGCCACCAAGGCCGATACCCAGCTGTCGCC 822  
Qy 916 TATCCCAATGAGGTCTTCCCAACACCTTTTGGTGAATGAAAGCCCTTCTATGATG 975  
Db 823 CACCCCAACGCGCGCTGCGGACCGCTTTTGGCGAGTACGACGAAACCCCGCGCGAGATG 882  
Qy 976 GCCAAGCACCTAAAGGATTTTGTCTATGATGGCTTTGGTCAATATATGTTGAGGATGCTGT 1035  
Db 883 GCGCGGTGTCGAAGAAATTCGCGCCAGCGCTTCTCAACATCATTTGGCGGTGCTGC 942  
Qy 1036 GGGTCAACACAGATCATATCAGGAAATTTCTGAGCTGTGAAATAATTTGAGCCTAGA 1095  
Db 943 GGCACCAACGCGCGGCGACATCAGGCGCATTTCCGAAAGCCGTGGCAAGATACAAGCCGCGC 1002  
Qy 1096 GTTCCACCTGCACTGCTTTTGAAGGACATATGTTACTGTCTGCTAGAGCCCTTCAGG 1155  
Db 1003 GAGATCCCGGAAATTCGCCAAGCCCTGCC-----GCCTGTGCGGCCCTGGAGCCGTTCAAC 1056  
Qy 1156 ATTGAGCGGTACACCAACTTTTGTAAACATTTGGAGAGCGCTGTAAATGTTGAGGATCAAGG 1215  
Db 1057 ATCGACCGCAGTCTGCTGCTCAACGCTGAGCGCCACCAACATCACCGTTCCGCC 1116  
Qy 1216 AAGTTTGCTAACTCATCATGCGCAGGAACTATGAAGAAGCCTTTGTGTGTCGCAAGTG 1275  
Db 1117 AAGTTGCGCAGCGTGTATCCGCGAAGAGAACTATACCGAAGCCCTGGAAGTCCGCTCGAG 1176  
Qy 1276 CAGGTGGAATGGAGCCAGGTGTTGGATGTCAACATGATGATGATGCGCATGCTAGATGGT 1335  
Db 1177 CAGGTGAGCGCGCGCCAGGTAAATCGACATCAACATGGAAGGAGGATGCTCGACTCC 1236  
Qy 1336 CCAAGTGCATGACAGATTTTGCAACTTAACTTCTCCGAGCCAGACATCGCAAAAGTA 1395  
Db 1237 CAGCGGCCATGCTCGGCTTCTCAACCTGATCGCGGTGAGCCGACATTTCCGCGTA 1296  
Qy 1396 CTTTGTGCTGATCTCTCAATTTTGTCTGATGAAAGTGGGTTAAAGTGTGCCAA 1455  
Db 1297 CCGATCATGATCGACTGTCTCAAGTGGAGATGATCGAAGCCGCGCTCAAGTGCATCCAG 1356  
Qy 1456 GGGAGTGCATTTGCAATGATTAATGATGATGAGAGAGAGAGAGACTTCTTGGAGAG 1515  
Db 1357 GGCAGGCGCATCGTCAACTCCATTTCCATGAAGGAGCGCTCGAGCAGTTCAACACCCAC 1416  
Qy 1516 GCCAGGAAGATTAAAAAGTATGAGAGCTGCTATGTTGCTCATGGCTTTTGTATGAAGAAGGA 1575  
Db 1417 GCCCGCTGTGCAAGCGCTATGGCGCTGCGGTGCTGATGGCTTCGAGCAAGTCGCG 1476  
Qy 1576 CAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGCGCTACCATCTGCTTGTG 1635  
Db 1477 CAGCGCGACACAGCGCGCGCAAGAAGAAATCTGCCAGCGCAGCTATGATCATCTGTGC 1536  
Qy 1636 AAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCCCTAATATCTTAACCATTTGGG 1695  
Db 1537 AACGAAGTGGGCTTTCCGCGCAGAGACATCATCTTCGACCCGCAACATCTTTTGGCGTGGC 1596  
Qy 1696 ACTGGAATGGAGGAACAACCTTGTATGCCATTAATTTATCCATGCAACAAAAGTCAAT 1755  
Db 1597 ACCGGTATCGAAGAGACAAACATCTACCGCTGACATTCATCGAGGCGCTGCGCTATATC 1656  
Qy 1756 AAAGAAACATTTACCTGGAGCGAGAAATAGTGGAGTCTTTCCAACTTGTCTTCTCTCTTC 1815  
Db 1657 CGCGACCACTGCGCCCGCGCTGAGCTCAGGTGCTGCTCAACAGTGTCTCTCGTTC 1716  
Qy 1816 CGAAGATGGAAGCCATTCGAGAGAGCAATGATCGGGTTTTCCTTTTACCATGCAATCAAG 1875  
Db 1717 CGTGGCAACAACCCGCGTGGTGAAGCGATCCACTCGGTGTTCTGTATACCGCATCCAG 1776  
Qy 1876 TCTGGCATGGACATGAGATGATGATGCTGGAACCTCCCTGTGTATGATGATCCAT 1935  
Db 1777 AACGGCCTGACCATGTTGGGTATCTGTCAGCGCGCGAGCTGGAGATCTTACGACGAGATCCCG 1836

Qy 1936 AAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGATATAAGACCTTGAGGCCACTGAG 1995  
Db 1837 GCCAGCTGCTGAAAAAGTTCGAGGACGTGTACTCAACGCGCAACCCCGCATGGCACCAGT 1896  
Qy 1996 AAGCTCTTACGTTATGCTCCAGACT---CAAGGCACAGGAGGGAAGAAAGTCATTTCAGACT 2052  
Db 1897 GCATGCTGCGCCATCGCGGACGACTACAGGGTGGCGGCGCAACCAAGGAAGTGGAAAC 1956  
Qy 2053 GATGAGTGGAGAAATGGCCCTGCGAAGAACGCTTGTAGTATGCTTGTGAAGGGCAAT 2112  
Db 1957 GAAAGTGGCGCTCGCTGCGGTGGAAGAAACGCTGAGCAGCGCTGTGCTCAAGGGCATC 2016  
Qy 2113 GAAAAACATATTATTGAGGATACCTGAGGAAGCAGGTTAAACCAAAAAAATATCCCGA 2172  
Db 2017 ACCCGCTTCATCTCGAAGACACCGAAG-----CATGCCGCCAGCAGTGCAGCGC 2067  
Qy 2173 CTTCTCAATATAATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTGGTGTATCTTTT 2232  
Db 2068 CCATCGAGTATCGAAGGCCCGCTGATGAACGCGATGAACGTTGCTGGGAGCTGTTC 2127  
Qy 2233 GGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAGTCAGCCCGGGTTATGAAGAAGCT 2292  
Db 2128 GCGCGGTAAATGTTTCTTCCGCGCAGGTGCTCAAGTCGCGCCGCGTCATGAAACAGGCC 2187  
Qy 2293 GTTGGCCACCTTATCCCTTTTCTATGGAAGAAAGAAAGAGAGAAACCAAGAGTCTTTAACGGC 2352  
Db 2188 GTTGTCTCACCTGATCCCGTTTCATCGAAGCGCAAAAGGCGACA----- 2230  
Qy 2353 ACAGTAGAAGAGGAGGCCCTTACAGGGGACCATCGTCTGGCCACTGTTTAAAGGCGAC 2412  
Db 2231 -----AGCGGAAGCCAAAGGGCAAGATCCTGATGGCCA CGGTGAAGGGCGAC 2277  
Qy 2413 GTGCACGACATAGCGAAGACATAGTTGGAGTAGTCTTGGCTGCAATAATTTCCGAGTT 2472  
Db 2278 GTGCACGACATCGCAAGAACATTTGCGCGTGTGCTGGCTGTACCGGATACGACATC 2337  
Qy 2473 ATTGATTTAGAGTCATGACTCCATGTGATGAAGATCTGAAGTGTCTTTGACCAACAA 2532  
Db 2338 GTCGACCTTGGCGTGTGTTACCGGGCGAAGATCTCGCAACCGGCGCGAGCAGAAG 2397  
Qy 2533 GCAGATATAATTGGCTGTGAGGACTCATCATCTTCCCTGGATGAATGATTTTGT 2592  
Db 2398 TGGACATCATCGGCTGTCCGCGCTGATTAACCGCTGCTGGACGAATATGGTCCAGTT 2457  
Qy 2593 GCCAAGGAATGAGAGATTAGCTATAAGGATTCATTTGATTGGAGAGCAACCACT 2652  
Db 2458 GCCCGGAATGACGCCAGGCTTCGAATCGCGCTGATGATCGTGGCGCACCAACC 2517  
Qy 2653 TCAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAG---TGCACTGTATTAATCAAT 2709  
Db 2518 TCCAAGCGCACACTGCAGTCAAGATCGAGCCCAAGTACAGCAACGACGCGGTCTATCTAC 2577  
Qy 2710 GTCTGGACCGCTTCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAATATCTAAG 2769  
Db 2578 GTCAACCGACTTTCGCGCGGCTCGCGCTGGCCACCCAGTTGTCTCAAGAAGTCAAG 2637  
Qy 2770 GATGAATACCTTTCAGGAAATCATGGAAGAAATATGAAGATATAGACAGGACCAATATGAG 2829  
Db 2638 CAGGCTTGTTCGAAGAAAGCCCGCTGGAATACGTGACGCTGTGAGCGTACCGCAAC 2697  
Qy 2830 TCTCTAAGGAGAGGAGATATCTTACCTTAAAGTCAAGCCAGAAAAAGTGGTTTCCAAATG 2889  
Db 2698 CGCAGCGCCCGCACCGAAACGCTCAGCTACGCCAGCGCATCGCGCAAGCGCGAGTAC 2757  
Qy 2890 GATTGGCTGTGAAACCTCACCAGTGAAGCCACGTTTATTGGGACCCAGGTCTTTGAA 2949  
Db 2758 GACTG---GGCGGCTACACGCGCGCGCTCTCTTCCCGCGCTCAAGGCTCTGGAG 2814  
Qy 2950 GACTATGACCTGCAGAGCTGTGGACTACATTGACTGGAAGCTTTCTTTGATGCTGG 3009  
Db 2815 GACATGATTTTGGCACCCCTGGCGGAGTACATGACTGGAGCGCGTCTTCTTCTCTCTGG 2874

Qy	3010	CAGCTCGGGGCAAGTACCGGAATCGAGGCTTCCCAAGATATTTAAACGACAAACAGTA	3069
Db	2875	GACCTGGCCGCGAAGTTTCCGCGC-----ATCCTCACCGACGAGGTGGTT	2919
Qy	3070	GGTGGAGAGCGCCAGGAAGGTTACAGATGATGCCACAAATATGCTGAAACACACTGATTAGT	3129
Db	2920	GGCGAGCCGCGCACCGCGCTGTACAAAGGATGCCGCGAGATGCTCGACAGCTGATCGAT	2979
Qy	3130	CAAAAGAAACTTCGGGGCCCGGGGTGTGGTTGGGTTCCTGGCCAGCACAGATATCCAAGAC	3189
Db	2980	GAGAAGCTGATCAGCGCTCGCGCGGTATTCGGCTTCTGGCCGGC-----CAAC	3027
Qy	3190	GACATTCACCTGTATCGCAGAGGCTGCTGCCCCAGGCTCGACAGGCCATAGCCACTTTC	3249
Db	3028	CAGGTGACGATGACGACATCGAAGTCTACGGCGAAGCGCCAGGCACATGGCCACCCTG	3087
Qy	3250	TATGGGTTAAGGCCAAAGGCTGGAAGGACTCTGCCAGCAGGAGCCATACTACTGCCTC	3309
Db	3088	CATCACCTCGCCAGCAGACCATCAAGCCGGAAGGCA-----AGCCCAACTGGTCGCTG	3141
Qy	3310	TCAGACTTCATCGCTCCCTTCCTTCGATTCGGCATCGGTGACTACCT---GGCGCTGTTGCC	3366
Db	3142	GCCGACTTCGTGCGCACCGAAAGACAGTGGGGCTTACCAGCTACGTAGGTGGTTTCATCACC	3201
Qy	3367	GTTCCCTGCTTTGGGGTAGAAGACTCAGCAAGGCCCTATGAGGATGATGCTGACGACTAC	3426
Db	3202	ACCGCCGGCATCGGTGCCGAGGAGTGGCCAAAGCTTACCAGGACAAGCGGCGACTAC	3261
Qy	3427	AGCAGCATCATGTGTCAAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAGAAGAGCTC	3486
Db	3262	AGCTCGATCATGTGTCAAGGCACTGGCCGACCGCCTGGCCGAAGCTGTGCCGAATGGCTG	3321
Qy	3487	CATGAAGAGTTCCCGAGAACTCTGGGCTTACTGTGCGAGTGAAGAGCTGGAGCTCGCA	3546
Db	3322	CACGAGAGGTGCGTTAAGAGCACTGGGGCTATGCCCCGAGCAGGACCTTGGACAAGCG	3381
Qy	3547	GACCTGCGAAGTTGCGGTCAAGGGGATCCGCCCGGCTCTGTGCTACCCGACGACGCC	3606
Db	3382	GCGCTGATCAAGGAACAGTACAGCGCATCCGCCCTCGCGCGGTATCCGGCTGCCCC	3441
Qy	3607	GACCACACCGAGACTCACCATGTGGAGACTCGAGACTCGAGACTC-----GAG	3651
Db	3442	GACCACACTGAAAAGGAAACTCTGTTCGCCCTGCTCATGTGACCGCGCATCGGTGAAACC	3501
Qy	3652	CAGTCTACGGCATTAGGTAAACAGAAATCATTAGCAATGGCACCTGCTTCAGCAGTCTCA	3711
Db	3502	GGGCCGAGCGGGGTACCTCACCGAGCACTTCGCGATGTTCCCGCGCGCGGTGAGC	3561
Qy	3712	GGCCTCTACTTCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAGATTTCCAGGAT	3771
Db	3562	GGCTGGTACTTCGGCCACCCGACGCGAAGTATTTTCCCGTAGCAAGGTGCAACAGGAC	3621
Qy	3772	CAGGTTCAGGATTATGCAATTGAGGAAGAAATATCTGTGGCTGAGGTGTGAAATGCTT	3831
Db	3622	CAGATCGAGCGTTACAGCGCACGCAAGGCCAGGATATCAGGTGACGAGCGCTGCGTG	3681
Qy	3832	GGACCCATTTTGGGATATGATA	3853
Db	3682	GCGCTTAACCTGGGGTATGACA	3703

## RESULT 9

US-10-282-122A-33284  
; Sequence 33284, Application US/10282122A  
; Publication No. US20040029129A1

GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel

```

1  / APPLICANT: Trawick, John
2  / APPLICANT: Carr, Grant
3  / APPLICANT: Yamamoto, Robert
4  / APPLICANT: Forsyth, R.
5  / APPLICANT: Xu, H.
6  / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
7  / FILE REFERENCE: ELITRA.034A
8  / CURRENT APPLICATION NUMBER: US/10/282,122A
9  / CURRENT FILING DATE: 2003-02-20
10 / PRIOR APPLICATION NUMBER: 60/191,078
11 / PRIOR FILING DATE: 2000-03-21
12 / PRIOR APPLICATION NUMBER: 60/206,848
13 / PRIOR FILING DATE: 2000-05-23
14 / PRIOR APPLICATION NUMBER: 60/207,727
15 / PRIOR FILING DATE: 2000-05-26
16 / PRIOR APPLICATION NUMBER: 60/230,335
17 / PRIOR FILING DATE: 2000-09-06
18 / PRIOR APPLICATION NUMBER: 60/230,347
19 / PRIOR FILING DATE: 2000-09-09
20 / PRIOR APPLICATION NUMBER: 60/242,578
21 / PRIOR FILING DATE: 2000-10-23
22 / PRIOR APPLICATION NUMBER: 60/253,625
23 / PRIOR FILING DATE: 2000-11-27
24 / PRIOR APPLICATION NUMBER: 60/257,931
25 / PRIOR FILING DATE: 2000-12-22
26 / PRIOR APPLICATION NUMBER: 60/267,636
27 / PRIOR FILING DATE: 2001-02-09
28 / PRIOR APPLICATION NUMBER: 60/269,308
29 / PRIOR FILING DATE: 2001-02-16
30 / Remaining Prior Application data removed - See File Wrapper or PALM.
31 / NUMBER OF SEQ ID NOS: 78614
32 / SOFTWARE: PatentIn version 3.1
33 / SEQ ID NO 33284
34 / LENGTH: 3717
35 / TYPE: DNA
36 / ORGANISM: Pseudomonas syringae
37 / US-10-282-122A-33284

```

Query Match	22.8%; Score 892.2; DB 7; Length 3717;
Best Local Similarity	55.2%; Pred. No. 2.9e-258;
Matches 2070; Conservative 0; Mismatches 1563; Indels 120; Gaps 12	
Qy	138 TCTCAGAAGAGGATTATGGTCTGGATGAGAGGATGGGACCACTATGATCCAGCGGGAGA 197
Db	42 TCTGAACAGAGATCTCGATTCTCGATGGCGCATGGCCACCATGATCCAGAGCTACAG 101
Qy	198 GCTAAACGAGAACACTTCCGAGGTCCAGGAATTTAAAGATCATGCCAGGCCGCTGAAGG 257
Db	102 GCTTGAAGAGAAAGACTATCGGGCCAAACGTTTTGCCACTGGCCCAAGTGACGTCAAGGG 161
Qy	258 CAACAAATGACATTTTAAAGTAAACTCAGCCCTGATGTCAATTTACCAAATCCATAAGGAATA 317
Db	162 CAACAAAGACTGCTGTGATCTGTACCCCTCCGACGTGATCGGCCCATCGAGAGCGGTA 221
Qy	318 CTTCTCGCTGGGGCAGANATCATATTGAAACAAATPACTTTTACGACGACATAGTATATGCCCA 377
Db	222 TCTGGATGCGGTGCCGACATCTTCGAAACCAACACCTTCAACGCGCACACAGGTGTCCCA 281
Qy	378 AGCTGACTATGGCCTTGAAACTTTGGCCTACCGATGAACATGTGCTCTGCAAGGAGTGGC 437
Db	282 GGCCGACTACCGCATGGATTCGATCTGTCTACAGCTGAACGTTCGAGGCGCCCGCTGGC 341
Qy	438 CAGAAAAGCTGCCGAGGAGGTAACTCTCCAGA---CAGGAATTAAGAGGTTTGTGGCAGG 494
Db	342 CCGTAAGGTTCGACAGCGCCAAAACCTTCGAAACCCCGGACAAAACCGCGCTTCGTTGCTGG 401
Qy	495 GGCTCTGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGGAAAGCCCGGATTA 554
Db	402 CGTGTCTGGCCCGACAGCCGCACTGTGTCTGTGTGCGCGGACGTCAACACCCCTGGGCTA 461
Qy	555 TAGGAACATCACTTTGATGAGCTTTGTTGAAGCATACCAAGAGCAGCGCCAAAGGACTTCT 614
Db	462 TCGCAACGTTCAGTTTCGACGAACCTGGTGGAGAACTATACCGAGGGCCACCAAGGCGCTGAT 521

QY	615	GGATGGCGGGTGTGATATCTTATCTCATATTGAAACTATTTTGTGATCTGCCAATGCCAAGGC	674
DB	522	CGAGGGCGGCGCGACCTGTATCTGTATCGAAGCAATCTTCGACACCTCAACGCCAAGGC	581
QY	675	AGCCTTGTGTGCACTCCAAAATCTTTTGTGAGGAGAAATATGCTCCCGGCTTACTTTAT	734
DB	582	GGCGATTTTCGCGGTGAGGGGTGTTTCGAAGAGTTGGGTTTTGAACTGCTTATCATGAT	641
QY	735	TTCAAGGACCATCGTTTGATAAAAAGTGGCGGACTCTTTTCGGGACAGACAGAGAGGAT	794
DB	642	TTCCGGCACCATCACCGACGCTTCGGCCGCACTCTGTCGGCCAGACTACCGAAGGTT	701
QY	795	TGTCATCAGCGTGTCTCATGAGGAAACAATCTGTGCAATTGGAATTAATTTGTGCTTTGGGTGC	854
DB	702	CTGGAATCTCGATCAGCCATCAAAACCGGTCTCCGTGGCTGAACTGCGCCCTGGGTGC	761
QY	855	AGCTGAGATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATCTCTCTG	914
DB	762	CAGTGAACTCGCGCTTTTACCTGCAAGAGCTGGCCACAAGGCCAACACTCATGTGTCTGC	821
QY	915	TTATCCCAATGCAAGGTCTCCCAACACTTTGGTGACTATGATGAAACGCTTCTATGAT	974
DB	822	GCACCCCAAGCCGGCTGCGCAACGGTTCCGCGAGTACGACGAATCTGCCGAGCCAGAC	881
QY	975	GGCCAAAGCACTTAAAGGATTTTGCTATGGAATGCTTGGTCAATATATAGTTGGAGATGCTG	1034
DB	882	TGCGAAGATCATCGAAGAAATTCGCTCAAGCGGTTTCTTGAAATGCTCGCGGCTGCTG	941
QY	1035	TGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAATAATGTGAAGCCTAG	1094
DB	942	CGGCACCACACTGAGCACATCAAGCGGATTCGCGAGGCATGTCTGGTTATGCACCCCG	1001
QY	1095	AGTTCCACCTGCGCACTGCTTTTGAAGACATATGTTACTGCTGTGGTCTAGAGCCCTTCAG	1154
DB	1002	TGAAATTCAGACATCCCAAAGCCTGCC-----GCCTGTGCGGCTTGGAACCGTTTAC	1055
QY	1155	GATTGGACCGGTACACCAACTTTGTTTAACTATGGAGAGCGCTGTAAATGTTGCAGGATCAAG	1214
DB	1056	CATCGATCGCACTGCTGCTTTCGTCAACGTCGCGAGCGCACCAACATCACCGGCTCCG	1115
QY	1215	GAATTTGTCTAACTCATCATATGGCAGAAACTATGAAGAACCTTGTGTGTTGCCAAAGT	1274
DB	1116	CCGCTTTGCGCGCTGATCCGTGAAGACAACTACACCGAAGCGCTGAGGTCGCCCTGCA	1175
QY	1275	GCAGTTCGAAATGGGAGCCAGGTGTTGGATGTCAATCGATGCGCATGCGCATGCTAGATGG	1334
DB	1176	ACAGGTCGAGCGCGGTGCCAGGTGATCGACATCAACATGGAGAGGGGATGCTCGACTC	1235
QY	1335	TCCAAGTGCATGACACAGATTTTGCAACTTAAATGCTTCCGAGCCAGACATCGCAAGGT	1394
DB	1236	GAAGAGGCCATGTGTCACTTCTCAACCTGATCGCAGGCGAGCCGGATATCTCTCGCGT	1295
QY	1395	ACCTTTGTGATCGACTCTCTCCAATTTTGTGTGATTAAGTGGGTAAAGTGTCTGCCA	1454
DB	1296	GCCGATCATGATCGACTCTCTCCAAGTGGGAAGTATCGAGGCGGTCTCAAGTGCATTCA	1355
QY	1455	AGGGAAGTGCATTGTCTAATAGCATTAGTCTGAAGGAAGGAGGACGACTTCTTGGAGAA	1514
DB	1356	GGGCAAGGGCATTCGTCAACTCGATCAGATGAAGGAAGGGGTGAGCAGTTTCATTCATCA	1415
QY	1515	GGCCAGGAAGATTAAAAGTATGAGGCTGTATGTTGGTGTATGCTTTTCATGAAGAAGG	1574
DB	1416	CGCTCGCTGTGCAAGCGCTACGGTCGGCAGTCTGTTGGTATGCGCTTCGACGAAACAGGG	1475
QY	1575	ACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGACCCGGGCTTACCATCTGCTTGT	1634
DB	1476	CCAGGCCGACACCGAAGCGCGAAGAGGAATCTGCAAGCGCTGTACGACATCTCTGGT	1535
QY	1635	GAAAAAATCTGGCTTTAAATCCAAATGACATTAATTTTGTGAACCTTAATATCTTAACCATTTG	1694
DB	1536	CAACGAAGTGGCTTTCCCGCGGAAGACATCATCTTCGAGCCGGAACATCTTTCCGCAATTCG	1595

1695	Qy	GACTGGNATGGAGACAACTTGTGATGCCATTAAATTTTATCCATGCAACAAAGTCAT	1755
1596	Db	CACCGGTATCGAAGACACAACTTACGCGGTGGACTTCATCAAAGCGCTGTGCCCTACAT	1655
1755	Qy	TAAAGAAACATTACCTGGAGCCAGATAAGTGGAGGTCTTTCCAACTGTGCTCTTCCTT	1814
1656	Db	CCCGATGAACTCCCGCAGCCCTGACCTCGGCGGCGGTGCNAAGTGTCTGCTTCTCGTT	1715
1815	Qy	CCGAGGAATGGAAGCCATTTCGAGAAGCAATGCAATGGGGTTTCTCTTTACCATGCAATCAA	1874
1716	Db	CCGCGGCAACAAACCCGGTACGCGAAGCCATTCACTCGGTGTTCTCTCTCCACGCGATCCG	1775
1875	Qy	GTCTGGCATGGACATGGAGATAGTGAATGCTCGAAACCTCCCTGTGTATGATGATATCCA	1934
1776	Db	CAATGGCCTGAGCATGGGCATGCTCAATCCGCGGCGAGCTGGAAATCTACGACAGATCCC	1835
1935	Qy	TAAGGAACTTCTGCAGCTCTGTGAAGATCTCATCTCGAAATAAAGACCTTGAGGCCACTGA	1994
1836	Db	GGCGAGCTGGCGCATGTGTGCAAGCGTGTGCTCAACCGTAAATCCCGAAGGCACCGA	1895
1995	Qy	GAAGCTCTTACGTTATGCT---CCAGACTCAAGGCAACAGGAGGAGAAAGTCAATCAGAC	2051
1896	Db	GGCGTTGCTGGCCATTGCCGACAAGTACAAGGCGCATGGCAGCGTCAAGGAGGCTGAAAC	1955
2052	Qy	TGATGAGTGGAGAAATGGCCCTCTCGAAGNAGCCCTTGATGTAGTATGCCCTTGTGAGGGCAT	2111
1956	Db	CGAAGAAATGGCGAGCTGGCCGGTCAACACAGCCCTGGAGCACGCGTGGTCAAGGGCAT	2015
2112	Qy	TGAAACAATATTATTGAGATACTCAGGAAGCCAGGTTAAACCAAAAMAAATATCCCGG	2171
2016	Db	CACATCCCATCTGTGCAGGACACCGAGAGTGCAGCTGGCTGGGC-----TTCACCGG	2066
2172	Qy	ACCTCTCAATATAAATTGAAGGACCCCTGATGAAATGGAATGAAATTTGTTGGTGATCTTTT	2231
2067	Db	GCCTATCGAGGTGATCGAAGGCCCGCTGATGTCGCGCATGAACGTGTGGTGCACCTGTT	2126
2232	Qy	TGAGCTGGAAAAATGTTTCTCACTCAGGTTATAAGTCAAGCCGGGTATGAGAAAGGC	2291
2127	Db	CGCGCGGGCAAGATGTTTCTGCGCAGGTGGTGAAGTCTGCAAGGGTGATGAAACAGGC	2186
2292	Qy	TGTTGGCCACTTATCCCTTTTCATGNAAGAAAGAGAGAAACCAAGAGTGCTTAAACGG	2351
2187	Db	GGTCGGCGACCTTATTCCCGTTCACTGAACCTGGAAGAAAGGCGCA-----	2230
2352	Qy	CACAGTAGAAGAAAGAGGACCCCTTACACAGGCAACATCGTCTGCCCACTGTTAAAGGCGA	2411
2231	Db	-----AGCCGAAGCCAAAGGCAAGATCCCTCATGCCACGCTCAAGGCGA	2276
2412	Qy	CGTGCAAGCATAGGCGAAGAACATAGTTGAGTAGTACTCTTGGCTGCAATTAATTTCCGAGT	2471
2277	Db	CGTTCAAGCATCGGCAAGAAACATCGTCGCGGTGGTGTGGGCTGTAAACGGCTATGACAT	2336
2472	Qy	TATTGATTAGGAGTCATGACTCCATGTGATAGATACTGAAGCTGCTCTTGACCAAA	2531
2337	Db	CGTCGATCTCGGTGTGATGGTGCCTCCGAGAGAATCTTCGAGGTGCCCGCGACGAGAA	2396
2532	Qy	AGCAGATATAATTGGCGCTGCAGACTCATCACTCTCTTCCCTCGATGAATGATTTTGT	2591
2397	Db	GTGCGACATCATCGGCTGTCCGGCTGATCAACCGCTCACTGGAAGAAATGSGTCAATGT	2456
2592	Qy	TGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCCATGTTGATTGGAGGAGCAACCAAC	2651
2457	Db	TGCCCGTGAATGCAGCGTCAGGACTTCCACCTGCCCTGATGATCGCGCGCGCACCAAC	2516
2652	Qy	TTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAG---TGCACCTGTAATCCA	2708
2517	Db	CTCCAAAGCGCACACGCGCGTCAAGATCGAGCCCTAAATACAGCAACGATGCACTGATCTA	2576
2709	Qy	TGTCCTGCAGCGCTCCAAGAGTGTGGTGGTGTGTTCCACGCTGTAGATGAAATCTTAA	2768
2577	Db	CGTCACCGACGCTCCCGACCGGTGGGGTCCGCAACCCAGTTCCTGCTCGAAGAGCTGAA	2636
2769	Qy	GGATGAATACTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACCACTATGA	2828

```
Db 2637 ACCGGCTTCATTGAAAAACCCGCCCTTGAAGTCGAGCGAAGCACGTCGGC 2696
Qy 2829 GTCTCTCAGGAGGAGATACCTTACCTTAAGTCAGCCGAGAAAGTGTTCCTCAAT 2888
Db 2697 GCGCAGTCGACGACCGAGCGCTGAGCTATGGCGCAGCGGTGGGAAAAAGCGAAGTT 2756
Qy 2889 GGATTTGGCTGTCAACCTCAACCCAGTGAAGCCACAGTTTATTCGGACCCAGGTCTTTGA 2948
Db 2757 CGACTGGAAAGTTACAC---CCCTGCAAGCCGACCTTCTACTGACCCGCGTCTTGCA 2813
Qy 2949 AGACTATGACCTGCAGAGCTGTGGACTACATTTGACTGGAGCCTTTCTTTGATGCTG 3008
Db 2814 GGACATCGACCTGAACGTGTGCTGCTACTCATCGACTGGAGCGCGTCTTCTTCAATTCCTG 2873
Qy 3009 GCAGCTCGGGGCAAGTACCCGAATCGAGGTTCCCAAGATATTTAAGCACAACAGT 3068
Db 2874 GGACCTGGCCGTTAAATACCCGCGC-----ATCCTGACCGACGAAGTGGT 2918
Qy 3069 AGGTGGAGAGCCAGGAAGGTCTACGATGATGCCCAACAATATGCTGAACACACTGATTAG 3128
Db 2819 TGGCGAAGCTCCACTGCATGTTTGAAGATGCCAGCAATGTTGGCAACTGATCGA 2978
Qy 3129 TCAAAAGAAACTCCGGGCCCGGGGTGTGGTGGTCTTGGCCAGCACAGATATCCAAAGA 3188
Db 2979 TGAATAACTGATCAGCGCGCGCGGTATTTCGGCTTCTGGCCGACCAACAGATCAACGA 3038
Qy 3189 CGACATTCACCTGTACGAGAGGCTGTGTGCCAGGCTGCAGAGCCATAGCACCTTT 3248
Db 3039 TGACGATC-----TGAGGTCTATGGCGAAGACGCGCAAGCGCTGGCGAAACT 3086
Qy 3249 CTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGCAGGACATACACTGCT 3308
Db 3087 GCATCACCCTGCCAGCAGACATCAAGCTCTGACGGCA-----AACGAAATTTCTCGCT 3140
Qy 3309 CTCAGACTTCATCGCTCCCTTGTCTTCTGGCATCCGTGACTACCT---GGCCCTGTTTGC 3365
Db 3141 GGCAGACTTCGTTGGCCCAAGGACAGCGCTGACCGACTACATTTGGTGGTTTCATCAC 3200
Qy 3366 GTTTCCTGCTTTGGGTTAGAGAGCTGAGCAAGCCCTATGAGGATGATGTGAGACTA 3425
Db 3201 CACCGCAGCATTTGGCGCGAGGAAGTGGCCAAAGCTTTATCAGGACAAACGCGCAGCTA 3260
Qy 3426 CAGCAGCATCATGTCAAGCGCTGGGGGACCGCTGGCAGAGCCCTTTGAGAGAGCT 3485
Db 3261 CAATCGATCATGTCAAGGCTTGGCCGACCGCTTGGCCGAGCGCTGCGCAAGCCTGCGCAATGGCT 3320
Qy 3486 CCATGAAAGAGTTCCCGAGAACTGTGGGCCCTACTGTGGCAGTGGAGCAGCTGGACGTGC 3545
Db 3321 GCACCAACAGGTGCGCAAGGATTACTGGGTTACGCGAAGACGAGCGCTGGACAAGA 3380
Qy 3546 AGACTCGGAAGTTGCGGTTCAAGGATCAAGGCGATCCGCCCGGCTCTGCTACCCAGCCAGCC 3605
Db 3381 GCGCGTATCAAGAGCAGTACATGGGCTATTCGCCCCGCTCCGGGCTATCCGGGCTGCC 3440
Qy 3606 CGACACACCGAGAGCTCACCATGT-----GGAGACT 3638
Db 3441 TGACCATACCGAAAGGGCACCTGTTGCGCCCTGCTCGACCCGCTGCGGAAGGCGAGCC 3500
Qy 3639 GCGAGACATCAGCAGCTCTACAGGCAATTAGGTTTAACAGAAATCATTTAGCAATGGCACCTGC 3698
Db 3501 CGAGCAGACCGCGGCAAAAGCGGAGTGTCTTGCAGCAGCAGCTACGCCATGTTCCCGGC 3560
Qy 3699 TTCAGCAGCTCAGCCCTCTACTTCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGAA 3758
Db 3561 GCGCGCAGTCAGCGGCTGTATTTTCGCTCATCTCATCCTCAAGCGCAGTACTTTGCAAGTCGGCAA 3620
Qy 3759 GATTTCCAGGATCAGGTTGAGGATTAATGCAATTGAGGAAGAACATATCTGTGGCTGAGGT 3818
Db 3621 GGTGCAAGGATCAGTTCGAAGCTACAGCAGCAGGAGGTGAGGACCTGAGGCTCAC 3680
Qy 3819 TGAGAAATGGCTTGGACCCATTTTGGGATATGA 3851
```

```
Db 3681 CGAACGCTGGCTGGCACCACCCAACTTGGCTACGA 3713

RESULT 10
US-10-282-122A-20596
; Sequence 20596, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20596
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-10-282-122A-20596

Query Match 22.6%; Score 883.8; DB 7; Length 3684;
Best Local Similarity 54.4%; Pred. No. 1e-255;
Matches 2043; Conservative 0; Mismatches 1612; Indels 102; Gaps 9;
```

```
Qy 120 GGATGAGATCAATGCCATTCTGCAGAAAGAGATTATGGTGTGGATGGAGGATGGGAC 179
Db 15 GGAACAACCTGCTGCGCAGTTAAATGAACGCTATTCTGCTGTGACGCGGTATGGGCAC 74
Qy 180 CATGATCCAGCGGAGAGCTTAACGAAGAACCTCCGAGGTCAGGAATTTAAAGATCA 239
Db 75 CATGATCCAGATTATCGACTGAACGAAGCGGATTTTCGTGGTGAACGCTTTGCCGACTG 134
Qy 240 TGCAGGCGCGCTGAAAGCGCAACATGACATTTTAAAGTATAACTCAGGCTGATGCTTTA 299
Db 135 GCCATGCGACTCAAGGCAACACGACCTGCTGCTACTCAGTAAACCGGAAGTATCGC 194
Qy 300 CCAATTCATGAAGTAATCTTGGCTGGGCGAGATATCATTTGAAACAATATCTTTAG 359
Db 195 CGCTATCCACAACGCCCTACTTTTGAAGCGGCGCGGATATCATCGAANACCAACCTTCAA 254
```

Db	1329	AGGTCTGAAGTGTATCCAGGCGCAAGAGCATTTGTTAACTCTATCTCGATGAAAGAGGCGCGT	1388
Qy	1497	GGACGACTCTTCGAGAGAGCCAGGAGAGATTAAAAAGTATGGAGCTGCTATGCTGCTGCAT	1556
Db	1389	CGATGCCCTTTATTCATCATCAGCGAAATTTTTCGCTCGCTACGGTGCAGAGTGGTGGTAAAT	1448
Qy	1557	GGCTTTTGTATGAAGAGGACAGGCAACAGAAAAACAGACACAAAAATCAGAGTGTGCACCCG	1616
Db	1449	GGCCTTTGACGAACAGGGAACAGGCCGATATCTCGCGCACGGAAATCGAGATTTTGGCGTGC	1508
Qy	1617	GGCCTACCATCTGTTGTGAAAAAACTGGGCTTTTATCCAAATGACATTTATTTTTCGACCC	1676
Db	1509	GGCGTACAAAATCCTCACCGAAGAGGTGGCTTCCGCGCAGAGATATCATCTTCGACCC	1568
Qy	1677	TAATATCCTAACCATTTGGGACTGGAATGGAGGAACACAATCTGTATGCCATTAAATTTTAT	1736
Db	1569	AAACATCTTCGCGGTGCAACTGGCATTTGAAGAGCACAACTACGCGCAGGACTTTTAT	1628
Qy	1737	CCATGCAACAAAAGTCAATTAAGAAAAATTTACCTCGAGCGCAGATAAGTGGAGGCTTTTC	1796
Db	1629	CGGCGGTGTGAAGACATCAACCGNACTGCGCGCACGGCTGATTTCCGGCGCGTATC	1688
Qy	1797	CAACTTGTCTTCTCTTCGAGGAATGGAGCCATTCGAGAAAGCAATGATGGGGTTTT	1856
Db	1689	TAACTTTCTTTCTCGTTCGTGCGCAACGATCCGGTGCAGAACCATTTACGCACTGTT	1748
Qy	1857	CCTTTACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACTCCC	1916
Db	1749	CCCTACTACGCTATTTGCAATGGCAATGGGATATGGGGATCGTCAACGCGCGGCAACTGGC	1808
Qy	1917	TGTGTATGATGATATCCATAAGGAACCTTCTCGAGCTCTGTGAAGATCTCATCTGGAATAA	1976
Db	1809	GATTTACGACGACTTACCCGCTGNACTGCGGACGGTTCGAAGATGTGATTTCTTAATCG	1868
Qy	1977	AGACCTTGAGGCGCATGAGAAAGTCTTTACGTTATGCCAGACTCAAGGC-----AC	2027
Db	1869	TCGCGACGATGGCACCGAGCGTTTACTTGGAGCTTGGCGAATAATATCGCGCAGCAAAAC	1928
Qy	2028	AGGAGGGAAGAAAGTCAATTCAGACTCATGATGTCGAGAAATGGCCCTGTGCAAGAACGCT	2087
Db	1929	CGACGACACCGCCAAACGCCCAGCAGCGGAGTGGCGCTCGTGGGAAGTGAATAAAGTCT	1988
Qy	2088	TGAGTAGCCCTTGTGAAGGGCAATCGAAAAACATATTTAGGATACTGAGGAAGCCAG	2147
Db	1989	GGATACTCGCTGTCBAAGGCATTAACCGNGTTTATCGACGAGATACCGAAGAGC---	2045
Qy	2148	GTTAAACCAAAAAATATCCCGACTCTCAATATAATTGAAGGACCCCTGATGAATGG	2207
Db	2046	-----CCGCCACGAGCTACGCGCCGATGAAGTGAATGAAGGCGCGTTGATGACCG	2099
Qy	2208	AATGAAAAATGTTGGTGATCTTTTTCGAGCTGGAATAATGTTTCTACCTCAGGTTATAA	2267
Db	2100	CATGAATGTGTCGCGGACCTGTTTGGCGAAGGGAATAATGTTCTCGCCACAGGTGTCAA	2159
Qy	2268	GTGAGCCCGGTTTATGAAGAGGCTGTGCGCACCTTATCCCTTTCATGAAAAAGAAAG	2327
Db	2160	ATCGGCGGCTCATGAAACAGGCGGTGGCTTACTTCGAACCGTTTATGAGCCAGCA-	2218
Qy	2328	AGAAGAAACAGAGTGTCTTAAACGCAAGTGAAGAGAGGAGCCCTTACAGGGCACCAT	2387
Db	2219	-----AAGAGCAGGCGCAAAACCAACGGCAAGAT.2246	
Qy	2388	CGTGCTGGCCACTGTTTAAAGGCGACGTGCACGACATAGGCAAGAACATAGTTGGAGTAGT	2447
Db	2247	GGTGATCGCCACCGTGAAGGCGGACGTCCACGACATCGGTAAAAATATCTGTTGGTGTGT	2306
Qy	2448	CCTTGGCTGCAATAATTTCCGAGTTATGATTTAGGAGTCACTGACTCCCATGTGATAAGAT	2507
Db	2307	GCTGCAATGTAAACATACGAAATGTGATCTCGGCGTTATGGTCCCTCGGAAAAAAT	2366
Qy	2508	ACTGAAAGCTGCTCTTGACCAACAAAGCAGATATAATTGGCTGTGAGGACTCATCACTCC	2567

Db 2367 TCTCCGTACCGCTAAAGAAAGTGAATGCTGATCTGATTGGCCCTTTTCGGGGCTTATCAAGCC 2426  
Qy |||||  
Db 2568 TTCCCTCGATGAATGATTTTGTTCGCAAGGAATGGAGAGATTAGCTATAAGATTCC 2627  
Qy |||||  
Db 2427 GTCCCTGAGCAGATGGTTAAACGTGGGGAAGAGATGGAGGCTCAGGGCTTCACTATTC 2486  
Qy |||||  
Db 2628 ATTGTGATTTGGAGGAGCAACCACTTCAAAACCCACACAGCAGTAAATAATAGTCTCCGAG 2687  
Qy |||||  
Db 2487 GTTACTGATTTGGCGGCGGACGACCTCAAAGCGCACACCGCGGTGAATATCGAGCAGAA 2546  
Qy |||||  
Db 2688 ATACAGTGCACTGTAAATCCATGTCCTGAGCGGTCCAAAGAGTGTGTGTGTTCCCA 2747  
Qy |||||  
Db 2547 CTACAGCGGCGGCGGAGCTGTATGTGCAGAAATGCTCGCGTACCGTGTGTGTGTGGGCGC 2606  
Qy |||||  
Db 2748 GCTCTTAGATGAATACTTAAGAGTGAATCTTTGAGGMAATCATGGAAGATATGAGA 2807  
Qy |||||  
Db 2607 GCTCGCTTCCGATACCCAGCGGTGATGATTTTGTGCTGCTGCTACCGCAAGGAGTACGAAC 2666  
Qy |||||  
Db 2808 TATTAGACAGGACCAATATGAGTCTCTCAAGGAGAGGATACCTTACCCCTTAAGTCAAGC 2867  
Qy |||||  
Db 2667 GGTAGCTATTACAGCAGGCGGCAAGAACCGGCACACACCCTGTCAGCTGGAGCGGC 2726  
Qy |||||  
Db 2868 CAGAAAAAGTGGTTTCAAAATGGATGGCTGTCTGAACCTCACCCAGTGAAGGCCACGTT 2927  
Qy |||||  
Db 2727 GCGCGATAACGATTTTCGCTTTTGACTGCGAGGCTTACACGCGCGCGGTGGCGCACCGTCT 2786  
Qy |||||  
Db 2928 TATTGGGACCCAGCTCTTTGAAGACTATGACCTGCAGAGCTGTGGACTACATTGACTG 2987  
Qy |||||  
Db 2787 CGCGGTGAGGAAGTCT-----GAAGCAGCATCGAAACGCTGCGTAAATACATCGACTG 2840  
Qy |||||  
Db 2988 GAAGCCTTTCTTGTATGCTGCGCAGCTCCGGGCGAAGTACCCGAATCAGGCTTCCCCAA 3047  
Qy |||||  
Db 2841 GACACCGTCTTTATGACTGGTCTGGTGGCGGGAAGTATCG-----CG 2885  
Qy |||||  
Db 3048 GATATTTAACCAAAACAGTAGTGGAGAGGCCAGGAGGCTTACAGATGATGCCCAA 3107  
Qy |||||  
Db 2886 CATTCGGAAGTGAAGTGTGGGCGTTGAGGCGCAGCGGCTGTTTAAAGACGCGCAACGA 2945  
Qy |||||  
Db 3108 TATGCTGAACACACTGATTAGTCAAAAGAACTCGGGCCCGGGTGTGGTGGTCTG 3167  
Qy |||||  
Db 2946 CATGCTGGATAAATTAAGCGCGGAGAAACCTGAACTCCGCGTGGCGGTGGGCGCTGTT 3005  
Qy |||||  
Db 3168 GCCAGCACAGATATCCAAGACGACATTCACCTGTACGACAGGCTGCTGTGCCCGCCAGGC 3227  
Qy |||||  
Db 3006 CCGCGCAACCGTGTGGCGATGACATTTGAATCTACCGTGACGAAACGCTACCCATG- 3064  
Qy |||||  
Db 3228 TGCAGAGCCCATAGCCATTTCTATGGTTTAAGGCAACAGCTGAGAGAGACTCTGCCAG 3287  
Qy |||||  
Db 3065 -----TGATCAACGTCAGCCACCATCTGCGTCAACAGACCGAAAAAAC 3107  
Qy |||||  
Db 3288 CACGAGGCATATCTACTGCTCTCAGACTTCATCGCTCCCTTGCATTCCTGCATCCGTTGA 3347  
Qy |||||  
Db 3108 AGGCTTGCCTAACTACTCTCTGCTGACTTCTGTTGCGCCGAAGCTTCTCTGTTAAGCAGA 3167  
Qy |||||  
Db 3348 CTACTGGGCTGTTTGGCGTGTGCTGCTGCTTTTGGGTAGAAGAG---CTGAGCAAGGCCTA 3404  
Qy |||||  
Db 3168 TTACATCGCGCATTTGCGTGACTGCGGGCTGGAAGAGACGACACTGGCTGATGCTT 3227  
Qy |||||  
Db 3405 TGAGATGATGGTGAAGACTACAGCAGCATCATGCTCAAGGCGCTGGGGGACCGGCTGGC 3464  
Qy |||||  
Db 3228 TGAAGCGCAGCAGATGATTAACAAATAATCATGTTGAAGCGCTTTCGCGACCGTTTACG 3287  
Qy |||||  
Db 3465 AGAGGCTTTGCAAGAGAGCTCCATGAAGAGTTTCGCGCAGAACTGTGTGGGCTACTGTGG 3524  
Qy |||||  
Db 3288 CGAAGCCTTTGGGAGTATCTCCATGAGCGGTGTGCGTAAAGTCTACTCGGGCTATGCGCC 3347  
Qy |||||  
Db 3525 CAGTGAGCAGCTGGAGCTCGCAGACTTCGGAAGGTTGCGGTACAAGGCATCCCGCGCGC 3584  
Qy |||||  
Db 3348 GAAAGAGAACTTCAGCAAGAGAGACTGATCCGCGAATACTACAGGCGCATCCGTCGCGC 3407  
Qy |||||  
Db 3585 TCCTGGCTACCCAGCCAGCCGACCAACCGAGAGAGCTCAACATGTGGAGACTCGCAGA 3644  
Qy |||||  
Db 3408 ACCGGCTTATCGGGCTCGCGGAAACATACGAAAAAAGCCCACTCTGGGAGCTGCTGGA 3467

Qy 3645 CATCGAGCAGCTCTACAGGCATTAGGTTAAACAGAAATCAATTAGCAATGCACTGCTTCAGC 3704  
Db |||||  
Qy 3468 AGTGAAGAAACACACTGGCATGAACCTCACAGAACTCTTCGCCATGTGGCCGCGTGCATC 3527  
Db |||||  
Qy 3705 AGTCTCAGGCTCTACTCTTCCCAATTTGAAGTCCAAATATTTGCTGTGGGGAAGATTTC 3764  
Db |||||  
Qy 3528 GGTTCGGGTGGTACTTTCAGCCACCGGACAGCAAGTACTACGCTGTAGCACAAAATTCA 3587  
Db |||||  
Qy 3765 CAAGGATCAGTTGAGGATTTATGATTGAGGAAGACATATCTGTGCTGAGGTTGAGAA 3824  
Db |||||  
Qy 3588 GCAGCATCAGTTGAAGATTTATGCCCGCCCTTAAAGGTATGAGCGCTTACCAGGTTGAGC 3647  
Db |||||  
Qy 3825 ATGCTTGGACCCCAATTTTGGGATATGATACAGACTAA 3861  
Db |||||  
Qy 3648 CTGCTGGCACCGAATCTGGGGTATGACGGGACTGA 3684  
Db |||||

## RESULT 11

US-10-282-122A-40847  
; Sequence 40847, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCES: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40847  
; LENGTH: 3681  
; TYPE: DNA  
; ORGANISM: Vibrio cholerae  
US-10-282-122A-40847

Query Match 22.5%; Score 883.2; DB 7; Length 3681;  
Best Local Similarity 54.7%; Pred. No. 1.6e-255;  
Matches 2032; Conservative 0; Mismatches 1583; Indels 99; Gaps 10;  
Qy 140 TGCAGAAAGAGATTATGTGTGCTGATGGGATGGGACCATGATCCAGCGGAGAGC 199



Db 38 TGAACACACGATCCTACTGATTTGATGGTATGCGTACCATGATTCAGAGTTATAAGT 97  
Qy 200 TAAACGGAAGAACACTTCGAGGTCAGGAATTTAAAGATCATGCCAGGCGCTGAAAGCCA 259  
Db 98 TACAAGAGGAGACTATCGCGGTGCAAGATTTGTGATTCGACATGTGATTTGAAGGAA 157  
Qy 260 ACAATGACATTTTAAAGTATAACTCAGCTGATGTATTTACCAATCATTAAGATACT 319  
Db 158 ATAAAGCACTCTTAGTGTCTTACTCAGCGCAATTTATTAAGAGATTCACCTCCGCTTACC 217  
Qy 320 TGCTGGCTGGGGCAGATATCATTAACAAATATCTTTAGCAGCACTAGTATTCGCCAAG 379  
Db 218 TTGAAGCGGGGCGGATATTTTGAACCAACACTTTAACTCAACCAAGATTCGCATGG 277  
Qy 380 CTGACTATGGCTTTGAACACTTGGCCCTACCGGATGAACATGTGCTCTGCGAGGTGCCCA 439  
Db 278 CAGACTATGACATGCAATCGCTCAGTGTGGAATTAACCTTTCGCGCGCTAAGCTTGCAC 337  
Qy 440 GAAAGCTGCGGAGGAGGTAA---CTCTCCAGACAGGAATTAAGAGTTTGTGCGAGGG 496  
Db 338 GTGAAGTCGCGGATGAGTGGAGCGCTAAAGATCAAGTTCGCCACACGCTATGTGCTGGT 397  
Qy 497 CTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGGAAGCGCGATATA 556  
Db 398 TGCTTGGGCCAACCAACCGTACTTGTCTTATTTCCGAGATGTGAACGATCCAGATTC 457  
Qy 557 GGAACATCACATTTGATGAGCTGTGTAAGCATACCAAGCAGCGCCAAAGACTTCTGG 616  
Db 458 GTAACGTCACTTTTGTGAGGCTGTGTGAGCTATTCGATCAACGCGGCTTTGATCA 517  
Qy 617 ATGCGGGGTGTATCTTACTCATTTGAACCTATTTTGTGATCTGCCAATGCCAAGCAG 676  
Db 518 AAGGTGGCAGCGATCTGATCTCATTTGAAACCACTTCGATACACTTAAAGCCAAAGCCT 577  
Qy 677 CTTTGTGTGACTCCAAATCTTTTGAAGAGAAATATGCTCCCGGCTATCTTTATTT 736  
Db 578 GTGCGTTTGGCGTGCATAGCGTATTTGAAGAGCTGGGCATCAGCTTACCTGTGATGATTT 637  
Qy 737 CAGGACGATGTTGATAAAGTGGCGGACTTTTCGGACAGACAGGAGGAGATTTG 796  
Db 638 CCGGACGATTTACGATGCGCTCTGGGAGACTCTGTGAGGACAGACAGGAGCTTCT 697  
Qy 797 TCATCAGCGTGTCTCATGAGAACCACTCTGCATTTGAATTAATTTGTCTTTGGGTGAG 856  
Db 698 ACAACGCTTGCCTCATGTAACGCGGATTTTGGTTGGCTTGAACCTGTGGGTAGGTCTG 757  
Qy 857 CTGAGATGAGACCTTTTATGGAATTAATGGAATAATGTAACAGCCTATGTCTCTGTT 916  
Db 758 ATGAGCTGCGCCAGTACGTGGAAGAGCTTTTACGCAATTTTCAAGATGTCTATGTTTCCGCG 817  
Qy 917 ATCCCAATGCAAGTCTTCCCAACACCTTTGTGACTATGTAAGAACGCTTCTATGATGG 976  
Db 818 ACCCAATGCGGACTGCGCAATGCTTTGTGATGATGATCTCTCTGCGAGGAATGG 877  
Qy 977 .CAAAGCACTTAAAGATTTTGTCTATGGAATGCTTTGGTCAATATATAGTTGAGGATGCTGTG 1036  
Db 878 CAGAACATATTTGCGGAATGGGACAAAGCTGCTTTTGAATTTGCTGGTGGTGTGCTGTG 937  
Qy 1037 GGTCAACACGATCATATCAGGGAATTTGCTGAGCTGTGGAATAATTTGTAAGCCTTAG 1096  
Db 938 GAACTACACCTGAGCATATCGCCGCCATTTGCCAAGCGCTGAGGGGTGTAACCAAGGG 997  
Qy 1097 TTCCACCTGCCACTTTCGAAGGACATATGTTACTGTCTGCTGCTAGGCGCTTCAGGA 1156  
Db 998 CTCTGC-----CAGATCGAAGTAGAATGTCTCTCTCGGGTTAGAGCGCTCAATA 1051  
Qy 1157 TTGACCGGTACACCAACTTTGTAAACATTGGAGAGCGCTGTAATTTGTGAGGATCAAGGA 1216  
Db 1052 TTGTCTCTGAACCTTTTGTAACTGGGGGAGCGTAACTAAAGTCAACCGCTTCTGCGC 1111  
Qy 1217 AGTTGTCTAACTCATCATGCGAAGAACTATGAAGAGCGCTTGTGTGTTGCCAAAGTGC 1276

Db 1112 GTTTTAAGCGTTTAAATTAAGAAGACCAATACGACGACGCTCGATGTGGCGGTGAGC 1171  
Qy 1277 AGGTGGAATPGGAGCCAGGTGTGTGATGTCAACATGATGATGCGATCTAGATGGTC 1336  
Db 1172 AAGTCGAATAACGCGCGCAGATCATTTGATATCAATGATGAAGGCATGTGGACCGC 1231  
Qy 1337 CAAGTGCATGACCAAGATTTTGCATTTTCTTCCGAGCCAGACATCGCAAGGTAC 1396  
Db 1232 AGCGGTGATGTTGCGCTTTTGAATCTATGCGCCTCTGAACAGAAATATCCAAAGTTC 1291  
Qy 1397 CTTTGTGCACTGCACTCTCCCAATTTTGTCTGATGTAAGCTGGGTAAAGTGTGCCAAG 1456  
Db 1292 CGGTGATGTCGACTCTCTTAATGGGAAGTCAATGAAGCGGCTCTGAATGCAATTCAGG 1351  
Qy 1457 GGAAGTGCATTTGTAATAGCAATTTAGTCTGAAGGAAGAGAGAGCACTTCTTGAGAGAG 1516  
Db 1352 GTAAAGGCATCGTCAACTCTATCTCTTAAGAAGAGGAAAGAGAAAGTTTATTTGCCAAG 1411  
Qy 1517 CAGGAGATTTAAAGATGAGGCTGCTATGTTGTGTCATGGCTTTTGTATGATGAAGAGAC 1576  
Db 1412 CCAATTTGGTGGCGCGCTACGGTGGCGGCTGATTTGTATGGCATTTTGAAGAGTGGGCC 1471  
Qy 1577 AGGCAACAGAAACAGACACAAATAACAGAGTGTGCACCCGGGCTACCATCTGCTGTGA 1636  
Db 1472 AAGCGGATACCGGTGAGCGCAATTAGAGATCTGTGCTCGGCTTACCATATTTTGGTGC 1531  
Qy 1637 AAAAACTGGGCTTTAAATCCAAATGACATTTATTTTGACCTTAATATCTTAACCAATGGGA 1696  
Db 1532 ATGAGGTGGGCTTTCCACCGGAAGATATTTATTTTGAACCGAACATCTTTGCTGTTCGA 1591  
Qy 1697 CTGGAATGGAAGAACCAACTTGTATGTCATTTAATTTTATTCATGCAACAAAGTCAATTA 1756  
Db 1592 CCGGAATTTGATGAGCAATAAATACGCACTGGAATTTTCAATATGCAAGTGGCGGACATTA 1651  
Qy 1757 AAGAAACATTACCTGGAGCCAGAAATAGTGGAGTCTTTTCCAACTGCTCTCTCTCTCC 1816  
Db 1652 AGGTGAGCTGCGCGATGCGATGATTTCTGGCGGTGTTCTAAGCTTTCTCTCTCTTCC 1711  
Qy 1817 GAGGAATGGAAGCCATTCGAGAAGCAATGATGGGGTTTTTCTTTTACCATGCAATCAAGT 1876  
Db 1712 GCGCAACAATATGTGGTGAAGCGATCCATGCTGTTTTTCTTTTATCACTGCTTCAAC 1771  
Qy 1877 CTGCGATGGAATGAGATGATGAAATGCTGGAACCTCTCTGTGATGATGATATCCATA 1936  
Db 1772 ACGGCATGGAATGGAATGTCGAACGCGAGGAGCTTGAATCTACGATAACCTTCCGC 1831  
Qy 1937 AGGAATCTTGCAGCTCTGTGAAGATCTCATCTGGAATTAAGACCTGAGGCCCATGGA 1996  
Db 1832 TGAACCTGCGTGAAGGAGTGGAAAGATGTGATCTCAATCGACGTAGCGTGGCAGGAAA 1891  
Qy 1997 AGCTCTTACCTTATGCGCAGACTCAAGGCACA-----GGAGGGAAGAAAGTCAATC 2047  
Db 1892 GACTGCTTGAGATCGCGAAGCGTATCGCGAAACAGTGTGGTAAAGAGAGGATGCTT 1951  
Qy 2048 AGACTGATGAGTGGAGAAATGCGCCCTGTGGAAGAACGCTTGTGATGATGCGCTTGTGAAG 2107  
Db 1952 CTGCATTTAGAGTGGCGCATGCGCTGTGCTGCTAAGCGCTTAGAGCAGCTTTTGTCAAAG 2011  
Qy 2108 GCATTTGAAGAACATATTTTGAAGTACTGAGGAGCCAGGTTAAACCAAAAAATATC 2167  
Db 2012 GCATACCGGAATTTATGCTCAAGACATCGAAGAGCACGTCAGCAAGCCAGTAA----- 2066  
Qy 2168 CCGACCTCTCAATATTAATTTGAAGGACCCCTGATGATGAAATGAAATTTGTTGGTGCATC 2227  
Db 2067 ----ACACTGGAAGTGAATTTGAAGGCGCTGATGATGATGATGATGATGATGATGATGAT 2122  
Qy 2228 TTTTGTGAGCTGGAAGAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTTATGAAGA 2287  
Db 2123 TGTTCGGGGAAGGAAATGTTCTCTACCGCAAGTCTGTAATAATCAGCGCGTGTCAATGAAC 2182  
Qy 2288 AGGTGTTGGCCACCTTATCCCTTTTATGAAAAAGAAAGAGAGAAACCAAGAGTGTCTTA 2347  
Db 2183 AAGCGGTGCGTATCTTTGAGCCTTTTCAATTAATGCGCAAAAAAGTGTGAGCA----- 2233

```
Qy 2348 ACGCCAGTAGAAGAAGAGGACCCCTTACAGGGCACCATCGTGTGCGCCACTGTAAAG 2407
Db 2234 -----CTTCAATGGTAAGATTTTCTGCGGACCGTAAG 2269
Qy 2408 GCGACGTGCACACATAGGCAAGCAATAGTTGGAGTAGTCTTTGGCTGCAATAATTCC 2467
Db 2270 GCGATGTGCATGACATTTGGTAAGAACAATTTTGGCGTGTCTGCTGCAAGTGTAAATTAACCTCG 2329
Qy 2468 GAGTATTGATTAGGAGTCACTGATCCATGTGATAGATACTGAAGAGCTGCTTTGACC 2527
Db 2330 AGATCATCGATCTTTGGTGTGATGTGCTTCGAGCAGATCCTCAAAAGTCGACGCGAGC 2389
Qy 2528 ACAAAGCAGATATAAATGGCTGTGAGGACTCATCTCCCTTCCCTGGATGAATGATTT 2587
Db 2390 AAAATGTCGATATCATCGGTCTCTCTGGCTTATCAAGCGCTCTTTGGATGAGATGTAC 2449
Qy 2588 TTGTTGCCAAGAAATGGAGAGATTAGCTATAAGGATTCCTATTTGATTGGAGGAGCAA 2647
Db 2450 ACGTGGCGAAAGAGATGGAGCGACAAGGGTTTGAAGTGCCTTTTGAATGGTGGGCAA 2509
Qy 2648 CCATTCAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACCTGTATCC 2707
Db 2510 CAACGTCTAAAGCGCATACCTCGGTGAAGATTGAACAGATATATCATGCGCTGTAGTG 2569
Qy 2708 ATGTCCTGGAGCGCTCCAAAGAGTGTGGTGTGTCTCCAGCTGTTAGATGAATCTAA 2767
Db 2570 ACGTGAATTAACGCGTCCGCGCGGTAGGGGTGTGCACATCATATTGTCTGATGAACAGC 2629
Qy 2768 AGGATGAATATCTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACCAATTATG 2827
Db 2630 GCCCGGATTTATCGAACGTTTGGATCTCGATTATAGCGCACGCGTATCAGCATGCTC 2689
Qy 2828 AGTCTCTCAAGGAGAGAGATACCTTAACCTTAAGTCAAGCCAGAAAAGTGGTTCCAAA 2887
Db 2690 GTAAAAACCCCAAAATCGCGCCAGTCACTGTAGAGCAGGACGCTGCTAAATAAGCGGCGC 2749
Qy 2888 TGGATTGCTGTCTGMACTCACCAGTGAAGCCACCGTTTATTTGGGACCCAGCTCTTG 2947
Db 2750 TGGATTGGGAAAATTAACAGCGCGCGC---CTCTGCGAAAACCGGGTGTGATGTGTTG 2806
Qy 2948 AAGACTATGACTGCAGAGCTGTGGGACTACATTGACTGAAGCCTTTCTTTGATGTCT 3007
Db 2807 AAAACATTTGGTTAGCCACACTACGTCTTATATCGATTGGACGCTTTTATGACTT 2866
Qy 3008 GGCAGCTCCGGGCAAGTACCGGAATCGAGCTTCCCAAGATTTTAAACACAAAACAG 3067
Db 2867 GGTGCTTTATGGCAATA-----CCCTGCCATTTTGGAGCATGAAGAGG 2911
Qy 3068 TAGGTGAGAGGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACACTGATTA 3127
Db 2912 TCGGTGAAGAGGCCAAAACGCTGTTTATGATGCCAATGCGCTTACTTGATAAAGTAGAGC 2971
Qy 3128 GTCAAAAGAAACTCCGGCCCGGGGTGTGGTTCTGCGCAGCACAGAGTATCCAAG 3187
Db 2972 GAGAAGGACTACTGAAGCCAGTGTATGTGCACTGTTCAGCAGCAGCGGTGGCG 3031
Qy 3188 ACGAATTTCACTGTATCGCAGAGGCTGTGTGCCCGAGGTGCGAGGCCCATAGCCACTT 3247
Db 3032 ATGACATTTGAGTGTACAGTGAATGAATCGCGTACGCAAGTGGCGCA-----TGTGC 3082
Qy 3248 TCTATGGGTAAAGCAAGCGCTGAGAGGACTCTGCCAGCAGGAGCCATCTACTGCC 3307
Db 3083 TGTACAACTTGGCTCAGCAGACTAGAA-----ACCGAAAAGGGGCCAACTACTGTT 3133
Qy 3308 TCTCAGACTTCACTGCTCCCTTGCATTCTGGCATCCGTGACTACTCGGGCTGTTTGGCG 3367
Db 3134 TGTCCGACTATGTTGCTCCGAAAGAGAGCGGTAAACGTTGATTTGGATTTGGCGCTTTGAG, 3193
Qy 3368 T---TGCCCTGCTTTGGGTGAGAAGAGCTGAGCAAGCCCTATGAGATGATGGTGCAGCT 3424
Db 3194 TAACTGTGGTCAATTGGTGAGCGAGCCTTTGGCGGATGCTTTATAAAGCTCAGGGTGTATGATT 3253
```

```
Qy 3425 ACAGCAGCATCATGTCTCAAGCGCTGGGGGACCGCTGGCAGAGAGCCCTTTGCAGAGAGC 3484
Db 3254 ACATGCGATCATGATCCAAAGCGGTAGCCGATCGTTTGGCGGAAGCCCTTTGGCGAATATC 3313
Qy 3485 TCCATGAAAAGAGTTGCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGAGCTGACGCTCG 3544
Db 3314 TGCATGAAAAAGTGCCTAAAGAGATTTGGCGGTTATGCGAGCGATGAATAATCTCTCCAATG 3373
Qy 3545 CAGACCTGCGAAGGTTGCGGTACAAGGGCATCCGCGCGCTCTGGCTTACCCAGCCAGC 3604
Db 3374 ATGACCTGATCCGTGAGCGTTTATCAGGGCATTCGACCCGCGGGGTATCCCGCGTGTCT 3433
Qy 3605 CCAGCACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGACAGCTCTACAGGCA 3664
Db 3434 CTGAGCATACGAGAAAGCGACTTTGTGGCAGATGCTAAATGTCGAAGAGACCATAGGTA 3493
Qy 3665 TTAGGTTAAACAGAAATCAATAGCAATGGCACTGCTGTTGAGAGCTCTAGGCTCTTCTTCT 3724
Db 3494 TGTCACTGACCAACAAGCTATGCGATGTGGCGCGGCTCTCGGTATCCGGTTGGTATTTCT 3553
Qy 3725 CCATTTTGAAGTCAAAATATTTGCTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGATT 3784
Db 3554 CGCATCCGATTTCTCGCTATTTTTCGGTATGCGGATCCAAACAGATCAACTGCACAGCT 3613
Qy 3785 ATGCAATTGAGAAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCCA 3838
Db 3614 ACGTGAAGCTTAAGGTTGGCGTTTGGAAAGCTGAATAAGTGGCTAGCGCCTA 3667
```

## RESULT 12

```
US-10-282-122A-39570
; Sequence 39570, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```

; SEQ ID NO 39570
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Salmo
US-10-282-122A-39

```

Query Match	22.1%;	Score 867.2;	DB 7;	Length 3684;
Best Local Similarity	54.4%;	Pred. No. 1.1e-250;		
Matches 208;	Conservative 0;	Mismatches 1598;	Indels 102;	Gaps 10;
QY	149	GGATTATGGTCTCGATGGAGGGATGGGACCATGATCCAGCGGGAGAGCTAAACGAAG	208	
DB	44	GTAATCTGGTGTGACGGCGGTATGGGACCAATGATCCAGAGCTATCGTCTACATGAAG	103	
QY	209	AACACTTCCGAGSTCAGGAATTTAAAGATCATGCGAGGCCGTGAAAGGCCAACAAATGACA	268	
DB	104	AAGATTTCCGGGGAGCGCTTTGCCGACTGGCGCCTCGACCTGAAAGGCCAACAAATGACC	163	
QY	269	TTTTAAGTATACTCAGCGTGAATGTCATTTACAAATCCATAGGAATACTTGTCTGGCTG	328	
DB	164	TGCTGGTCCCTCAGCAAGCGGAGGTGATCGCCGCTATCCACAACGCGCTACTTTGAGCGTG	223	
QY	329	GGGCAGATATCATTTAGAAACAAATACTTTTAGCAGCACTAGTATTGCCAACAGTCACTATG	388	
DB	224	GCGGGATATCATCGAAACCAACCTTTTAACCTCGACAACCAATTCGATGGCGGATTACC	283	
QY	389	GCCTTGAACTCTGGCCTACCGGATGAACTATGTCCTTCGCAAGGAGTGGCCAGAAAAAGCTG	448	
DB	284	GGATGGAATCCCTGTCGCGGAAATTAACATATGCGCGGCGCAAACTGGCGCGCCCTGCG	343	
QY	449	CCGAGGAGTAACTCTCCAGA--CAGGATTAAGAGTTTGTGGCAGGGGCTCTGGGTC	505	
DB	344	CCGATGAATGGACGGCGCGAACACCGAAAAACCAACGCTTTGTTGCGGGCGTGTCTGGTC	403	
QY	506	CGACTAATAAGACACTCTGTGTGCCCATCTGTGGAAGGCCCGGATATATAGGAACATCA	565	
DB	404	CAACTAACCGCACGGCCTCCATTTGCGCGGAGCTCAACGACCCCGCGTTTCGTAATATCA	463	
QY	566	CATTTGATAGCTTTGTTGAAGCATACCAAGACGAGGCCAAAGGACTTCTGATGGCGGG	625	
DB	464	CTTTGATAGCTGGTGGCGGCTTACCGTGAATCCACAAAGCCCTGGTGGAGCGCGGG	523	
QY	626	TTGATATCTTACTCATTTGAACTATTTTGTATCTGCGCAATGCCAAGCGAGCCCTTGTGTTG	685	
DB	524	CGGACCTGATCCTGATTGAACTGTCTTCGACACCTTCAAGCCCAAGCGCGGTGTGTTG	583	
QY	686	CACCTCCAAATCTTTTGGAGAGAAATATGCTCCCGCGCCTATCTTTATTTTCAGGACGA	745	
DB	584	CGGTGAAGAAGAGCTTTGAAGCGCTGGCGTTGATCTGCCGATCATGATTTCCGGCACCA	643	
QY	746	TCGTTGATAAAGTGGCGGACTCTTTCGGACACAGACGAGAGAGGGAATTTGTCTATCAGCG	805	
DB	644	TCACCGACGCGCTCTGGCGGTACGCTTTCGGGCCAGACGACGGAAGCCCTTTTATTAACTCGC	703	
QY	806	TGCTCTATGAGAACCACTCTGCAATCGAATTAATTTGCTTTTCGGTGCAGCTCAGATGA	865	
DB	704	TGCGCCACGCGGAGCGCTCACTTTTGCCCTTAAGTGGCCCTGGGGCCAGATGACTGC	763	
QY	866	GACCTTTTATTTGAATAATTTGGAATAATGTAACAACGCTTATGTCTCTGTTTATCCCAATG	925	
DB	764	GCCAGTACGTCACGAACCTGTCGCGGATTGCGGAATGCTAGCTCACCGCGCACCCGAACG	823	
QY	926	CAGTCTTTCCCAACACCTTTGGTGACTATGATGAACCGCTTCTATGATGGCCGAAGCACC	985	
DB	824	CCGCGCTTCCGAACCGCTTTCCGCGAGTACGACCTCGACCGCCGACCAATGGCGAAAAAGA	883	
QY	986	TAAAGGATTTTGCTATGGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACAC	1045	
DB	884	TTCCGGAATGGCGGGAAGCGGCTTCTGGAATATCGTTGGCGGCTGTCTGGCCACCGC	943	
QY	1046	CAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAAATTTGTAAGCCTTAGAGTTCCACCTG	1105	
DB	944	CGGAGCATATTTGCGCGGATGAGCGCGCTTTGCGGTTTCTCCGCGCGCAGCTGCGG	1003	

Qy	1106	CGACTGCTTTTGAAGGACATATGTTACTGTCTGGTCTAGAGCCCTTCAGATTTGGACCGT	1116
Db	1004	ATATCCCGGTGGCTGCG-----GCCTTTTCCGGCCCTGGAGCGCGTGAACAATTTGGTGACG	1057
Qy	1166	ACACCAACTTTGTAAACATTTGGAGAGCGCTGTAATGTTGCAGGATCAAGGAAGTTTGCTA	1225
Db	1058	ATAGCTGTTTGTCAACGTCGGCGAAGCTAATACGTCACCGCTCGGCCAAATTTAAAC	1117
Qy	1226	AACTCATCATGGCAGGAAAATATGAAGAAGCCTTTGTGTGTGTCACAAAGTGCAGGTGGA	1285
Db	1118	GCTTGATCAAGAAGAGAAATACAGCGAGCGCTGGATGTCGCCGCTGCAGCAGCTCGAAA	1177
Qy	1286	TGGAGCCCAAGGTGTGGATGTCAACATGATGATGGCATGCTAGATGTTCCAAGTGC	1345
Db	1178	CGCGCGCGCAGATTAATTTGATATCAATATGATGAGGGGATGCTCGAGCGCGAAGCGCGA	1237
Qy	1346	TGACCGAGATTTTGCACACTTAATTTGCTTCCGAGCCAGACATCGCAAAAGTACCTTTGTGCA	1405
Db	1238	TGTGTGGTTCCTCAGCCTGATTTGCCGCTGAGCCGGACATTTGCCGTGTACCAATCATGA	1297
Qy	1406	TGCACTCCTCCAAATTTTGTGTGATTAAGACTCGGTAAAAGTGTCTGCCAAGGGAAGTGC	1465
Db	1298	TTGACTCTCCAAATGGAGGTTATCGAAAAAGGCTGAGTGCATTCAGGGTAAAGGCA	1357
Qy	1466	TTGTCAANTAGCATTTAGTCTGAAGAAAGGAGAGACGACTTCTTTGAGAAAGGCCAGGAAGA	1525
Db	1358	TCGTCAACTCTATTTGCGATGAAGAGGGGTGGAAGCCTTTATTCATCATGCGAAGTTGC	1417
Qy	1526	TTAAAAGTATGAGCTGTATGGTGTGATGCTTTTGATGAAGAGGACGCAACAG	1585
Db	1418	TACGTGCTACGGCGCGCAGTGTGTGATGCTTTTGATGAGCAGGGGCAAGCCGAC	1477
Qy	1586	AAACAGACACAAAAATCAGAGTGTGCACCGCGCTACCATCTGCTGTGCAAAAAACTGG	1645
Db	1478	CCGCGAACGTAATTCGAGATTTGCCGCGCGCTTACAAAAATTTGCTCGAAGAGTAG	1537
Qy	1646	GCTTTAATCCAAATGACATTAATTTTGAACCTTAATATCTTAAACATTTGGGACTGG	1705
Db	1538	GCTTTCCGCGGGAAGACATCATCTTGCACCGAATATCTTGCCTGCGCACCGGTATTG	1597
Qy	1706	AGGAACAAACTTTGTATGCCATTAATTTTATCCATGCACAAAAAGTCATTAAGAAACAT	1765
Db	1598	AAGAGCAACAACACTACGCGCAGGACTTTATCGCGCTGTGGAAGACATCAAAACGCGAGC	1657
Qy	1766	TACTGGAGCCAGAAATAAGTGGAGGCTTTTCCAACTTGTCTCTCTCCGAGGAATGG	1825
Db	1658	TGCCGACGCGCTGATCTCGCGCGGGGTCTTAACGTGTCTTCTGTTTCGCGGCAACG	1717
Qy	1826	AAGCCATTCGAGAAAGCAATGCATGGGTTTTCCTTTACATGCAATCAAGTCTGGCATGG	1885
Db	1718	ACCGGTACGTGAGGCTATCCACGCGTATTCCTCTACTACGCCATCCGCAACGGCATGG	1777
Qy	1886	ACATGGAGATAGTGAAATGCTTGGAAAACTCCTGCTGTATGATATATCCATAAGGAATTC	1945
Db	1778	ACATGGGCATCTCAACCGCGGGCACTGCGGATTTATGACAACTGCTGCGCAACTGC	1837
Qy	1946	TGCAGCTCTGAGATCTCATCTGGAATAAAGACCTGAGGCCACTGAGAAGCTCTTAC	2005
Db	1838	CGGATGAGTTGAAGATGTCAITCTTAACCGTCCGATACGGCACCAGCGGTTTGTCTGG	1897
Qy	2006	GTTATGCCAGACTCAAGGCACAGGGGGAAGAAAGTCAIT-----CAGACTGATG	2056
Db	1998	ATTTGGCGGAGAAATATCGCGGCAGCAAAACCGACGAGGCTGCCAGTCCACGAGCGG	1957
Qy	2057	AGTGGAGAAATGCGCCTGTGCAAGAAAGCGCTTGAGTATGCCCTTGTGAAGGGCAATTGAAA	2116
Db	1958	AATGGCGTAGCTGGACGCTGAAAAAGCGTCTCGAATACTCGCTGGTGGTGAAGGCAITACCG	2017
Qy	2117	AACATATTATGAGGATACCTGAGGAGGCCAGGTTAAACCAAAAAAATATCCCGACCTC	2176
Db	2018	AGTTTATCGAACAGGATACCAAGAGAGCC-----CGTCAGAGCGCGCGCGCCGA	2068

Qy 2177 TCAATATAATTGAAGGACCCCTGATGAATGAATGAATAAATTTGTTGGTATCTTTTGGAG 2236  
Db 2069 TTGAGGTGATTGAAGGGCCGCTGATGGACGCATGAACGTGTCCGGACCTGTTCCGGC 2128  
Qy 2237 CTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTATGAAGAAGCTGTTG 2296  
Db 2129 AAGGGAATAATGTTCTCGCGCAGGTGGTGAATCCGCTCGGTGATGAACAAGCGGTG 2188  
Qy 2297 GCCACCTTATCCCTTTTCATGTAAAGAAAGAGAGAAACAGAGTGCTTTAAACGCACAG 2356  
Db 2189 CCTACCTGGAGCCGTTATTGAAGCCAGCAAGAAAAA ----- 2226  
Qy 2357 TAGAAGAAGAGGACCCCTTACAGAGGCACCATCGTGTGGCCACCTGTTAAAGCGCACGTGC 2416  
Db 2227 -----GGCTCCAGCAACGGCAGATGGTGAATGTCACCGTGAAGGGCGATGTGC 2275  
Qy 2417 ACACATAGGCAAGAACATAGTTGGAGTAGTCTCTGGCTGCAATAATTTCCGAGTTATTG 2476  
Db 2276 ACACATTTGGCAAGAACATTTGTCCGCGTGTGCTGCAATGCAACACTACGAAATCGTGC 2335  
Qy 2477 ATTTAGGAGTCATGACTCCATGTGATAGATACCTGAAGAGCTGCTTGAACCAAAAGCAG 2536  
Db 2336 ATCTTGGCGTATGTGTCAGCGGAGAAATCTCAGAACGGCGGTGAAGTGAATGCGC 2395  
Qy 2537 ATATAATTGGCTGTGAGGACTCATCTCTCTCCCTGGATGAATGAATTTTGTGCCA 2596  
Db 2396 ATCTGATTGGTCTTTCCGGGCTTATCACCCGCTGCTGGACGAAATGCTCAACGTTGCCA 2455  
Qy 2597 AGGAAATGGAGAGATAGCTATAGGATTCATTTGTTGATTTGGAGGAGCAACCACTTCAA 2656  
Db 2456 AAGAGATGGAGCGTCAGGGCTTTACTATCCCGCTACTGATCCGGCGGCGCAACCACTTCCA 2515  
Qy 2657 AAACCCACACAGAGTTAAATAGCTCCGAGATACAGTGCACTGTAACTCATCTCTCGG 2716  
Db 2516 AAGCGCATACCGCGGTGAAATCGAGCAGAACTACAGCGGTCCGACGCTTACGTGCGAGA 2575  
Qy 2717 ACGGCTCCAAAGATGTGGTGTGTTCCAGCTGTTAGATGAATAATCTAAAGATGAAT 2776  
Db 2576 ATGCTTCGCGTACCGTGGCGGTGGTGGCGGCTACTCTCCGACACCCAGCGTGTACT 2635  
Qy 2777 ACTTTGAGGAATATGAAAGAAATGATGAATATAGACAGGACCATATGAGTCTCTCA 2836  
Db 2636 TTGTGCGCCGTACCGCGAAAGAGTACGAAACCGTGCCTATTTCAGCACGCCCGCAAAAAC 2695  
Qy 2837 AGGAGAGAGATACCTTACCTTAACTCAAGCCAGAAAGGTTCTCCAAATGATTTGGC 2896  
Db 2696 CGCGCACCCCGCGGTCAAGTGGAGCGGCGGATATGATCTGGCATTTGATTGG 2755  
Qy 2897 TGTCTGAACCTCACCCAGTGAAGCCCGTATTATTTGGACCCAGGTCTTTGAAGACTATG 2956  
Db 2756 AACGCTACACCCCGCGGTAGCCACCGTCTGGGCGTCCAGAGGT -----GGAGGCCA 2809  
Qy 2957 ACCTGCAAGCTGTGTGACTACATTTGATGAGTGAAGCCCTTTTGAATGTTGGAGCTCC 3016  
Db 2810 GCATCGAAACCTGTGCGCACTACATCGACTGGACCGCTTTCTTTATGACCTGTGCTGCTG 2869  
Qy 3017 GGGCAAGTACCGGAATCGAGCTTCCCAAGATATTTAAGCAGAAACAGTGTGGAG 3076  
Db 2870 CCGCAAAATACCCG -----CGCATTTCTGGAAGATGAGGTGGCGGCTT 2914  
Qy 3077 AGGCGAGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTGATTAGTCAAAAGA 3136  
Db 2915 AGGCGCAGCTGTGTTTAAAGACGCCATGATATGCTGATATAACTGAGCGCCGAGAAC 2974  
Qy 3137 AACTCCGGGCGGGGTGTGGTGTGCTGGCCAGCAGAGTATCCAAAGACACATTC 3196  
Db 2975 TGTGAAATCCCGGTGGCGTGTGGGCTGTGTTCCCGCGAACCGTGTGGGTGACACATCG 3034  
Qy 3197 ACCTGTACGAGAGCTGTGTGCCCCAGGCTGCAGAGCCCATAGCCACTTTCTATGGT 3256  
Db 3035 AAATCTATCGGACAAAACCCGCTACTCTATGTTCTGACCGTCA -----GCCACCAACC 3085  
Qy 3257 TAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGGCCATACCTACTGCTCTCAGACT 3316

Db 3086 TGGCCACGACAGACCGAAGAAAGTTGGTTTGTCTA -----ACTACTGTCTGGCGATT 3136  
Qy 3317 TCATCGCTCCCTTGCATTCCTGCGCATCGTGTACCTACCTGGGCTGTTTGGCTTGCCTGC- 3375  
Db 3137 TTGTCCGCGCGAAACTGAGCGGCAAGCGGACTACATCGGTGTCTTTCGCGGTGACCGGCG 3196  
Qy 3376 --TTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACGACTACAGCAGCA 3433  
Db 3197 GTCTGAAGGAGATGCGCTGGCGGACGCTTTCGAGCGCAACACGACGACTTATACAGA 3256  
Qy 3434 TCATGTTCAAGGCGCTGGGGACCGGCTGCGAGAGGCTTTTGCAGAAAGCTCCATGAAA 3493  
Db 3257 TCATGTTGAAAGCGATTTGCCGACCGTCTGGCGAAGCGTTTTCGCGAGTATCTGCAATGAC 3316  
Qy 3494 GAGTTCCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCTGGAGCTGCGAGACCTGC 3553  
Db 3317 GTGTACGTAAAGTTTACTGGGATATGCGCGGAAACGAGAGCCCTGAGTAAACGAGAAATTAA 3376  
Qy 3554 GAAGGTTTGGGTACAAAGGCGATCCGCCGCTCTGCTACCCAGCAGCCGACCCAGCACA 3613  
Db 3377 TCCGGAATACTACAGGGGATTCGCCCGGCGCGGTTATCTGCTGCGCCGGAACATA 3436  
Qy 3614 CCGAAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGCAATTAGGTTAA 3673  
Db 3437 CCGAAAAAGGCACTATCTGCGAGCTACTGATGTGCAAAAAACACACCGGATGAAGCTCA 3496  
Qy 3674 CAGATCATTTAGCAATGGCACTGCTTACGACGCTCTCAGGCTCTACTTCTCCAAATTTGA 3733  
Db 3497 CCGAATCTTTGCGCATGTGGCTGGCGCTCGGTCTCGGCTTACTTTCAGCCATCCTG 3556  
Qy 3734 AGTCCAAATATTTGCTGTGGGAAAGATTTTCAAGGATCAGGTTGAGGATTTATGCAATTGA 3793  
Db 3557 AGACAAATACTTTCGCGGTAGCGAGATCCAAACGATCAGGTGACAGATTTATGCTTCC 3616  
Qy 3794 GGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTTTGGGATATGATA 3853  
Db 3617 GTAAAGGAATGAGCGTTTGAAGACGTTGAGCGGTGGCTCGCGCGAACCTTGGGTTACGATG 3676  
Qy 3854 CAGACTAA 3861  
Db 3677 CGGACTGA 3684

## RESULT 13

US-10-282-122A-24011

; Sequence 24011, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zvekind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA 034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

;; PRIOR APPLICATION NUMBER: 60/230,347  
;; PRIOR FILING DATE: 2000-09-09  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/267,636  
;; PRIOR FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 78614  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 24011  
;; LENGTH: 3681  
;; TYPE: DNA  
;; ORGANISM: Klebsiella pneumoniae  
US-10-282-122A-24011

Query Match 21.9%; Score 859; DB 7; Length 3681;  
Best Local Similarity 54.4%; Pred. No. 3.4e-248;  
Matches 208; Conservative 0; Mismatches 1595; Indels 102; Gaps 11;  
Qy 149 GGATTATGGTCTCGATGGAGGATGGGACCATGATCCAGCGGGAAGCTAAACGAAG 208  
Db |||||  
Qy 44 GTATTCTGGTCTCGATGGGGAATTTAAAGATCATGCGAGGCTGAAAGGCAACGACG 103  
Db |||||  
Qy 209 AACACTTCCGAGGTCAGGAATTTAAAGATCATGCGAGGCTGAAAGGCAACGACG 268  
Db |||||  
Qy 104 AGGACTTTCGGGTGAGGCTTCTGCTGCGGCTGCGACCTGAAAGGCAACGACG 163  
Db |||||  
Qy 269 TTTTAAATATTAACCTCAGCTGATGATTTTAAAGATCATGCGAGGCTGAAAGGCAACGACG 328  
Db |||||  
Qy 164 TGCTGGTCTCAGCAACCGAGGATCGCGAGATCCAGGACCTATTTTCGAGGCG 223  
Db |||||  
Qy 329 GGGCAGATATCATTTGAAACAAATCTTTTACGACACTGATATTTGCCAAGCTGACTATG 388  
Db |||||  
Qy 224 GCGCGATATCATTTGAGACCAACACCTTTTAACTCGACGCCATCGCGATGGCGGATPAC 283  
Db |||||  
Qy 389 GCCTTGAACACTTGGCCTACCGGATGAACATGCTCTGCGAGGATGCGCCAG ---AAAAG 445  
Db |||||  
Qy 284 AGATGGAATCCCTGCGCGGAGATCAACTTTGCGCGGAGCTGGCGCGCGCCAGCG 343  
Db |||||  
Qy 446 CTGCGGAGGAGTAACTCTCCAGACAGGAATTAAGAGGTTTGGCAGGCGCTCTGGGTC 505  
Db |||||  
Qy 344 CCGAGGCTGGAACGCGCGGACGCGGGAACCGCGCTATGTCGCGGCGTGTGGCC 403  
Db |||||  
Qy 506 CGACTAATAAGACTCTCTGTGTCCCATCTGTGGAAGGCGCGGATTTAGGAACATCA 565  
Db |||||  
Qy 404 CGACCAACCGCACCGCCTCTATCTCGCTGATGTGAACGATCCGCGCTTCGGTAATTA 463  
Db |||||  
Qy 566 CATTTGATGAGCTTTTGAAGCATACCAAGAGGCGCAAGGACTTCTGATGGCGGG 625  
Db |||||  
Qy 464 CTTTGAACAGCTGTGTGCGCGCTTATCGGAAATCACCCGCGCGCTGTGTGGAAGCGGCG 523  
Db |||||  
Qy 626 TTGATATCTTACTCATTTGAACATATTTTGAATCTGCAATGCAAGCGAGGCTTTGTTG 685  
Db |||||  
Qy 524 TGGATCTGATCTGATTTGAACGCTGTTGACACCTTGAACGCTTAAAGCGGCACTATG 583  
Db |||||  
Qy 686 CACTTCAAAATCTTTTGAAGGAATATGCTCCCGGCTTATCTTTATTTTCAGGACGA 745  
Db |||||  
Qy 584 CGGTGAAGAGGAGCTGGAGCGTTAGCGGTTGACCTGCGCTGATGATCTCCGGACCA 643  
Db |||||  
Qy 746 TCGTTGATTAAGTGGGCGACTCTTTCCGACAGACAGAGAGGATTTGTCATACGCG 805  
Db |||||  
Qy 644 TTACCGATGCTCTGCGCGCAGCTTTTCCGACAGACCGAGGCTTTTATAACTCGC 703  
Db |||||  
Qy 806 TGTCTCTGAGAACCACTCTGCAATGGAATTTAAATTTGCTTTGGTGGAGCTGAGATGA 865  
Db |||||  
Qy 704 TGCCTCAGCGGAAGCCTTATCGTTTGTGTGAATTAAGACCCCTGAGGCCACTGAGAGCTCTTAC 763  
Db |||||

Qy 866 GACCTTTTATTGAATAATTGGAAATGTACAAAGCCTATGTCCTCTCTGTTATCCCAATG 925  
Db |||||  
Qy 764 GGAGTACGTTCAGGAGCTGTCCGCAATTCGCGAATGCTACGTACCGCGCACCCGGAACG 823  
Db |||||  
Qy 926 CAGGTCTTCCCAACACACTTTGGTGACTATGATGAAAGCGCTTCTATGATGGCAAGCACC 985  
Db |||||  
Qy 824 CCGGCTTGCCCAACGCTTTTCGGCGAATACGATCTGGACGCCGACACCATGCGACGCAAA 883  
Db |||||  
Qy 986 TAAAGGATTTTGTCTATGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACAC 1045  
Db |||||  
Qy 884 TTCCGCAATGGCGGAAGCGGTTTCTCAATATCGTTGGCGCTGCTGCGGTACCAACC 943  
Db |||||  
Qy 1046 CAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAAATTTGAAGCTCTAGAGTTCCACCTG 1105  
Db |||||  
Qy 944 CCGAGCATATCGCGCCCATAGAGCCGCGGTAGCGGCGCTGCGCGCGCTGCTGTTGCCG 1003  
Db |||||  
Qy 1106 CCACTGCTTTTGAAGGACATATGTTACTGTCTGCTCTAGAGCCCTTCAGGATTTGGACCGT 1165  
Db |||||  
Qy 1004 AAATTGGGTGGCTCGCGC-----CTCGCGCGCTGGAGCGTTGAAATATCGCGGACG 1057  
Db |||||  
Qy 1166 ACACCAACTTTTGTAAATTTGAGAGGCGGTGTAATGTTGAGGATCAAGAAAGTTTGCTA 1225  
Db |||||  
Qy 1058 ATAGCCTGTTTCTGTTAACTGCGTGAACGAATTAACGTCACCGCTCGCGCAAAATTTAAGC 1117  
Db |||||  
Qy 1226 AACTCATATGCGCAGGAATTAAGAGGCTTGTGTGCTGTCGCAAAAGTACAGGTGGA 1285  
Db |||||  
Qy 1118 GGCTGATTAAGAAGAGAGTACAGGAAAGCGTGGACGTTGCGCGCTGACAGGTGGAAA 1177  
Db |||||  
Qy 1286 TGGGAGCCAGGTGTTGGATGTCAATGATGATGCAATGATGCTAGTAGTGGTCCAAAGTCA 1345  
Db |||||  
Qy 1178 GCGCGCGGAGATTCGATATCAATCAATGACGAGGAGTGTCTGACGCGGAGCGGCGA 1237  
Db |||||  
Qy 1346 TGACCAAGATTTTGCATTAATTTGCTTCCGAGCGACACATCGCAAAAGGTACCTTTGTGCA 1405  
Db |||||  
Qy 1238 TGGTGGTTCCTTAACTGATTTGCGCGGAGCGGACATTTGCCGCTGCCGATCATGA 1297  
Db |||||  
Qy 1406 TCGACTCTCTCAATTTTGTGATTTGAAGCTGGTTAAAGTGTCTGCAAGGAGTGTGCA 1465  
Db |||||  
Qy 1298 TTGACTCTCTCAAGTGGGAAGTCAITTTGAAGAGGCGCTGAAATGTAATTCAGGCGCA 1357  
Db |||||  
Qy 1466 TTGCTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCGACTTTCTTGGAGAGCGCAGGA 1525  
Db |||||  
Qy 1358 TCGTTAACTCGATTCGATGAAGAGGCGTCAATCTTTATCCACCATTCGAAACTGG 1417  
Db |||||  
Qy 1526 TTAAGAAATGATGAGCTGCTATGTTGCTCATGCTTTTGAATGAAGAGGAGCAGGCAACAG 1585  
Db |||||  
Qy 1418 TTCCGCGCTACGCGCGCGGTGTGTGATGSCCTTCGACGAGTGTGTGTCAGCGGACA 1477  
Db |||||  
Qy 1586 AAACAGACACAAATCAGAGTGTGACCGCGGCTTACCATCTGCTTGTGAAAAAACTGG 1645  
Db |||||  
Qy 1478 CCGCGCAACGTAATAATTGAGATTTGCGCTGCGCGCTTACAAATCTCACCAGAGAGTGC 1537  
Db |||||  
Qy 1646 GCTTTAATCCAAATGACATTTTGGACCTTAATCTTAAACATTTGGAGCTGGGATGG 1705  
Db |||||  
Qy 1538 GCTTCCCGCGGAAGATATTTCTCGACCCCAATATCTTTGCGCGCTGCGCACCGGATCG 1597  
Db |||||  
Qy 1706 AGGAACAACATTTGATGCTTAAATTTTATCATGCAACAAAGTCAITTTAAAGAAACAT 1765  
Db |||||  
Qy 1598 AAGAGATTAACAATCAGCAGGACTTTATCGCGCGCTGCGAGGACATCAACGCGAGC 1657  
Db |||||  
Qy 1766 TACCTGGAGCAGAAATTAAGTGGAGGCTTTTCCAACTGTGCTCTCTCTCCCGAGGAATGG 1825  
Db |||||  
Qy 1658 TGCGCATGCGCTAACTCTCGCGGCGGTCTCCAACTGCTCTCTCTCTCTCTCTCTCTCT 1717  
Db |||||  
Qy 1826 AAGCCATTCGAGAGCAATGATGCGGTTTCTTTTACCATGCAATCAAGTCTGGCATGG 1885  
Db |||||  
Qy 1718 ACCCGTACGTGAGGCGATCCACGCGTATTCCTCTACTACGCGATCCCGCAACGCGATGG 1777  
Db |||||  
Qy 1886 ACATGAGATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATCCATTAAGGAATTC 1945  
Db |||||  
Qy 1778 ACATGGGATTCGTAACGCGGTGACGTGCGCATCTATGATGACCTCCCGGGGAGCTGC 1837  
Db |||||  
Qy 1946 TGCAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGAGCTCTTAC 2005  
Db |||||



1838 GCAGCGGGTGAAGAGCGTGAATCTCAACCGCCGCGATGACAGTACCGAAACGGTTACTTG 1897  
2006 GTTATGCTCCAGACTCAAGGCACAGAGGAGAAAGTCAATTCAGACTGAT-----G 2056  
1898 AGCTGGCGGAAATAATATCGCGGACGAAAGCGGACGACGGCGCAATGCCAGACGCGG 1957  
2057 AGTGAGAAATGGCCCTGTGGAAGAACGCTTGAAGTATGCCCTTGTGAAGGGCAATGAAA 2116  
1958 AGTGCGCACCTGGGAAGTGAAGAAACGCTCGAATATTCGCTGTTAAGGGCAATACCG 2017  
2117 AACATATTTATGAGGATCTGAGAACGACGTTTAAACCAAAAAAATATCCCGACCTC 2176  
2018 AATTATCGAACGACACCGAAGAC-----GCGTCAGCAGCGCCCGCCGCA 2068  
2177 TCATATATTTGAAGGACCCCTGATGAATGAAGTGAAGTATGTTGTTGATCTTTTGGAG 2236  
2069 TTGAGGTCAATTGAAGGGCCGCTGATGACGCGCATGAACGTTGTCGGCGATCTGTTGCGG 2128  
2237 CTGAAAAATGTTTCTACCTCAGGTATTAAGTCAAGCCCGGTTATGAAGAGGCTGTTG 2296  
2129 AAGCAAAATGTTCTGCGCGAGTGTGAATATCCGCCCGTGTATTAAGAACGCGGTGG 2188  
2297 GCCACCTTATCCCTTTCATGGAAGAAAGAGAGAAACACAGAGTGTCTTAACGGCACAG 2356  
2189 CCTACCTTGAGCCGTTTATTGAAGCCAGTA----- 2218  
2357 TAGAAGAAAGGACCCCTTACGAGGGCACCATGCTGCTGGCCACTGTTAAAGGCGACCTGC 2416  
2219 ---AAGAGCGAGGCTCCAGTAACCGGCAAAATGGTGTATGCGCCAGGTAAAGGCGACCTC 2275  
2417 ACACATAGGCAAGAACATAGTTCGAGTAGTCTTGGCTGCAATAATTCGAGTTATTG 2476  
2276 ATGACATTTGGCAAAACATCTGCGCGTGTGCTGCGAGTGCATTAATCTAGAAATCATCG 2335  
2477 ATTTAGGAGTCATGATCCATGTGATGAAGATPACTGAAGCTGCTTTTGACCAACAAAGCAG 2536  
2336 ATCTTGGCGTATGCTCCCGCGGATAAATTTCTCAAAACGCGCCAAAGAGTAAACCGG 2395  
2537 ATATAATTTGGCTCTCAGGACTCATCTCTCTCCCTGGATGAATGATTTTGTGCGCA 2596  
2396 ATCTAATTTGCTTTTTCGGCGGTGATTAACCCATCGCTGGACGAGATGTTAATGTGCGA 2455  
2597 AGAAATTTGAGAGATTAGCTATAAGGATTCATTTGTTGATTGGAGGAGCAACACTTCAA 2656  
2456 AAGAGATGGAGCGCGAGGCTTTACCATCCGCTGCTGTATAGGCGGCGGACCACTCGA 2515  
2657 AAACCCACACAGAGTTAAATATAGCTCCGAGATACAGTGCACCTGTATTCATGCTCTGG 2716  
2516 AAGCGCATACGGCGTGAATAATCGACGAGAACTACAGCGGCCCCACCGTTTACGTACAGA 2575  
2717 ACGGCTCAAGAGTGTGTGTGTGTTCCAGCTGTGTATAGTGAATAATCTAAGGATGAAT 2776  
2576 ACGGCTCGCGACCGTCCGCGTGTGTGTGCGCTCTCTCGGACACGCGAGCGACGAGT 2635  
2777 ACTTTGAGGAATCATGGAAGATATGAAGATATTAGACAGGACCATATGAGTCTCTCA 2836  
2636 TTGTGCGCGCACCGGCAAGAGTATGAACGTTGCGTATCCAGCATGGGCGTAAAAAC 2695  
2837 AGGAGAGAGATATCTTACCCCTTAAGTCAAGCGAGAAAGTGGTTTCCAAATGATTTGGC 2896  
2696 CGCGTACGCGCGGTGACCTTCTCCCGCGCGCGGAAACGATCTGCCCTTTGACTGGG 2755  
2897 TGTCTGAACCTCACCCAGTGAAGCCCGAGTTTATTTGGGACCCAGGCTTTTGAAGACTATG 2956  
2756 AGAGCTACACCCCGCGGTG-----CTCATGCTCTGGGAGTACAGACGTTGGAAGCCAGT 2811  
2957 ACCTGCAAGAGCTGTGTGAGTACATGACTGGAAGCTTTCTTTGATGCTGCGAGCTCC 3016  
2812 ATC--GAGACTCTGGGCACTACATGCACTGAGCGCGTCTTCTCATGACCTGTGCTGGTGG 2869  
3017 GGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAACGACAAACAGTAGTGGAG 3076

2870 CGGTAATATATCCG-----CGCATTTCTGGAAGATGAAGTGGTAGGGAAG 2914  
3077 AGCCAGGAAGGTTCTAGATGATGCCACATAATGCTGAACACACATGATTTAGTCAAAAGA 3136  
2915 AGCGAGCGCCTGTTCAAAGACGCTAAACGAGTCTCTGGATAAGCTTAGCGCGAGAGAAA 2974  
3137 AACTCCGGGCGCGGGTGTGTTGGGTTCTGGCCAGCACAGAGTATCCAAAGACGACATTC 3196  
2975 CCCTGAACCCGCGCGGTGTGGTCTGTTCCCGGCCAACCGGTGGGCGACGATATTG 3034  
3197 ACCTGTACGACAGAGGCTGTGTGCCCGCAGGCTGAGAGCCCATAGCCACTTTTATGGGT 3256  
3035 AATCTATCTCGATGAACCCGCAACCCACG-----TGCTGACGCTGAGGCATCATC 3085  
3257 TAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGCCATACTACTGCTCTCAGACT 3316  
3086 TCGCCAGCAGACGAGAGAGGTTGGCTTCGCCA-----ACTACTGTCTCGCCGAT 3336  
3317 TCATCGCTCCCTTTGCAATTTCTGGCATCCGTACTACCTGCGGCCCTGTTTGGCGCTGCTGCT 3376  
3137 TTGTCGCGCCGAGACTGAGCGGCAAGCGGACTACATCGCGGCTTCGCGGTGACCGCG 3196  
3377 TTGGGTTAGAAGA---GCTGAGCAAGCCCTATGAGGATGATGTGAGACTACAGCAGCA 3433  
3197 GCCTGGAAGAGGACGCGCTGCGGACGCTTATGAAGCGCAGCATGATGACTATAACAAGA 3256  
3434 TCATGCTCAAGCGCTGGGGACGCGCTGCGAGAGGCTTTGAGAGAGCTCCATGAAA 3493  
3257 TCATGATTAAAGCCATCGCGATCGGCTGCGGAGGCTTTTGGCGAGTATCTCGATGAGA 3316  
3494 GAGTTCCGCGAGAACTGTGGCCCTACTGTGGCAGTGAAGCTGAGCTGCGAGACCTGC 3553  
3317 AAGTGGCAAGTTTACTGGGATATGCGGCCAATGAGAACCTCAGCAATGAGGAGCTTA 3376  
3554 GAAGTTGCGGTCAAGGGCATCCGCGCGGCTCTGCTGCTACCCAGCGCAGCCGACACA 3613  
3377 TTCGAGAACTACCGGGGATCCGACCGCGCGGGATATCCGGCTGTCCGAGCACA 3436  
3614 CCGAGAGCTCACCATCTGGAGACTCGCAGACATCGAGCAGTCTACAGGCAATAGGTAA 3673  
3437 CCAGAAAGGGGACCATCTGSCAGCTACTGACACTGGAAGCCCATACCGCATGAAGTCA 3496  
3674 CAGAACTATTAGCAATGGCACCTGCTTTCAGCAGTCTCAGGCTCTACTTCTCCAATTTGA 3733  
3497 CCGAGCTTTTCGCAATGTGCGCAGGCGCATCGGTTTCGGCTGCTACTTCAGCCATCCG 3556  
3734 AGTCCAAATTTTGTGTGGGAAAGATTTCCAAAGGATCAGGTTGAGGATATGCAATGA 3793  
3557 ACAGTAAGTACTTTCGCGCTGCGCAGATCCAGCGCATCAGGTGGAAGATTACGCGCTGC 3616  
3794 GGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTGAGCCCATTTTGGGATATGATA 3853  
3617 GTAAAGGAATGACCGCGCGGAGGTGAGCGCTGTTAGCGCTTAATCTGGGCTATGATG 3676  
3854 CAGAC 3858  
3677 CGAC 3681

## RESULT 14

US-10-282-122A-38912  
; Sequence 38912, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangau  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant







## RESULT 15

US-10-282-122A-41615  
; Sequence 41615, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haeelbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282.122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed -- See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41615  
; LENGTH: 3696  
; TYPE: DNA  
; ORGANISM: Yersinia pestis  
US-10-282-122A-41615

Query Match 21.8%; Score 854.4; DB 7; Length 3696;  
Best Local Similarity 54.7%; Pred. No. 8.4e-247;  
Matches 2042; Conservative 0; Mismatches 1591; Indels 103; Gaps 13;  
  
Qy 140 TGCAGAGAGATTATGTGCTGGATGGAGGATGGGACCAATGATCCAGCGGAGAGC 199  
Db 50 TGGCGCAACGATTTTGTGCTGATGCTGTATGGGACCAATGATCCAGAGCTACCGAT 109  
  
Qy 200 TAAAGGAAGAACACTTCGAGGTGAGGATTTAAAGATCATGCCAGCGCTGAAAGCA 259  
Db 110 TAGAAGAAGCAGATTACCGAGCGCGGCTTTGCTGACTGGGCAAGTGACCTAAAGGTA 169  
  
Qy 260 ACAATGACATTTTAAATGATATCAAGCTGATGTCATTTACCAATCCATAAGCAATAC 319  
Db 170 ACAATGACCTTTGGTGTCTTCAAGCCAGAGATCATCGCGGATCCATAATGCCCTACC 228  
  
Qy 320 TGCTGGCTGGGCGAGATATCAATGAAACAAATACCTTTTAGCAGCACTAGTATTGCCCAAG 379  
Db 230 TTGAAGCGGTGCTGACATTCGAAACCAATACCTTTTAACTCTACCTCCATTCCGATGG 289  
  
Qy 380 CTGACTATGGCCCTTGAACACTTTGGCCTACCGGATGAACATGTGCTCTGCGAGAGTGCCCA 439  
Db 1423

Db 290 CGGACTATCAGATGGCGCTCACTGTGGCTGAAATTAACATATGAAGCCGCCGTCTGCCCC 349  
Qy 440 GAAAAGCTGCCGAGGA---GGTAACTCTCCAGACAGGAATTAAGAGGTTTGTGCGAGGG 496  
Db 350 GTATCTGTGCGGATGAATGAGCGCGCGCAGCAAAACACACGTTATGTTGCCGGGG 409  
Qy 497 CTCTGGGTCCGACTTAAGACACTCTCTGTGTCCTCATCTGTGGAAGGCGCGGATATA 556  
Db 410 TGCTGGGGCCAAACCAACCGTACTGCTCCATCTCCCTAAAGTTAATGACCCGCGTTC 469  
Qy 557 GGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCCCAAGGACTTCTGG 616  
Db 470 GTAACGTCACTTTTGTATCAGTTGTCGAGCGCTATCTGTAGTCTACCGCGCTGATTG 529  
Qy 617 ATGCGGGGTGATATCTTACTCATTTGAAACTATTTTGTATGATGTCGCAATGCCAAGGAG 676  
Db 530 AAGGGGGTGTGATTTAATATGATCGAAACCGGCTCTCGATACGTTGAATCGAAGCGCG 589  
Qy 677 CCTTGTGTCACCTCCAAATCTTTTGGAGGAAATATGCTCCCGGCTATCTTTATTT 736  
Db 590 CGACCTTTGCGGTGAAGCGAATTCGAAGTATGGCGGTGCTGTTGCCGCTGATGATTT 649  
Qy 737 CAGGAGCATCGTTGATAAAAGTGGCGGACTCTTCCGACACACAGAGGAGGATTTG 796  
Db 650 CCGCACCATTACCGATGCTCCGCGTACGCTGCTGTCTCAACACGACGAGACTTTT 709  
Qy 797 TCATCAGCGTCTCATGAGAACCACTCTGCAATGGATTAATTTGCTTTGGGTGCGAG 856  
Db 710 ATAACTCGCTGCGCAATGTCAACCGCTGCTGTTGCTGTAACCTGTGCGCTGGCGCG 769  
Qy 857 CTGAGATGAGACCTTTTATTTGAATAATTCGAAATATGTAACAGCCTATGTCCTCTGTT 916  
Db 770 ATGAACCTGCGCCAATATGTTGCTGAGTTATCGGCTATCTCTGAATATTAATGTCAGT 829  
Qy 917 ATCCCAATGCACTCTTCCCAACACCTTGTGCTGACTATGATGAACCCCTTCTATGATGG 976  
Db 830 ACCGAAATGCGGTTTGGCGAATGCTTTTGGTGAATGATCTGGAAGCCAAAGATGG 889  
Qy 977 CCAAGCACCTAAAGGATTTTGTCTATGGATGGCTTGGTCAATATATAGTTGGAGGATGCTGTG 1036  
Db 890 CTGAGCAGATTGGCGAGTGGCGACGTCGCGGATTTTGAATATTTGTCGGCGTTGCTGCG 949  
Qy 1037 GGTCAACACAGATCATATCAGGGAATTCCTGAGCTGTGAAAAATTTGAAGCTTAGAG 1096  
Db 950 GTACCAACACCCCGCATATTTGCGCGATGTCACGCTGTGCTGGGTGAGAGCCCTTCAAG 1009  
Qy 1097 TTCCACCTGCACTGCTTTTGAAGGACATATGTTACTGTCTGGTCTAGAGCCCTTCAAG 1156  
Db 1010 CGTTGCTGATATTCGGGTTGCTGCGGTTTGG-----CGGGCTGGAGCCTGACTA 1063  
Qy 1157 TTGACCGTACACCAACTTTTGTAAATTCGAGAGCGCTGTAAATTTGTCAGGATCAAAGGA 1216  
Db 1064 TCGATGCCATACCTGTTTGTGACGTCGCTGAGCGGACCAACGTCACGGTTCCGAC 1123  
Qy 1217 AGTTTGTAACTCATGCGCAGGAACTATGAAGAGCCTTGTGTGTTGCCAAAGTGC 1276  
Db 1124 GCTTTAAACGCTGATAAAAGAGAGAGATGACGAGAGGCGCTGGATGTTGCCCGCAAC 1183  
Qy 1277 AGTGGAAATGGAGCCAGCTGTGATGTCATCACTGATGATGTCATGATGCTAGATGGTC 1336  
Db 1184 AGTTCGAAACGCGCGCGAGATCATCGATATCAACATGGAAGGAGGATGTTGGATGCCG 1243  
Qy 1337 CAAGTGAATGACAGATTTTGAACCTTAATTCCTCCGAGCAGACATCGCAAGGTAC 1396  
Db 1244 AAGGGCCATGTTGGCTTCTTAATCTGATTCGCGGGAGCTGATATTGCCAGGCTGC 1303  
Qy 1397 CTTTGTGCACTGACTCCTCCAAATTTTGTGCTGATGAAGCTGGTTAAAGTGTGCCAAG 1456  
Db 1304 CGATTATGATCGACTCTCAAGGTGGATGTGATTTGAAAAGGCGCTGAAATGATTTCAAG 1363  
Qy 1457 GGAAGTGCATTTCAATAGCATTTAGTCTGAAGGAGGAGGAGGACTTCTTGGAGAGG 1516  
Db 1364 GCAAAGCATCGTGAACCTCCATTTTCGATGAAGAGGCGGTTGATGCTTTTATTCATCACG 1423

QY 1517 CCAGGAAGATTAAAGATTTAGAGCTGCTATGGTGGTCAATGGCTTTTGTATGAAGAGGAC 1576  
DB 1424 CCAACTGGTACGGCGTTATGGTCGCCCATGGTGGTCAATGGCTTTGTATGAAGAGGAC 1483  
QY 1577 AGGCAACAGAAACAGACACAAATAATCAGAGTGTGCACCCGGGCGCTACCATCTGCTGTGA 1636  
DB 1484 AAGCCGATACCGCGCAGCAAAATCGAGATTTGCGCGCGTGGCTATAAAATATTGACGG 1543  
QY 1637 AAAAATCGGGCTTTAATTCAAATGACATTTATTTTGGACCCCTAAATATCTTAACCATTTGGGA 1696  
DB 1544 AAACCGTGGTTTCCCAACAGAGATATTTCTTCGATTCGAAATATTTTGGCGGTAGCGA 1603  
QY 1697 CTGGAATGGAGAACACAACTTTGATGCAATTAATTTTATCCATTCGCAACAAAGTCAATTA 1756  
DB 1604 CGGGATTTGAAGCATTAACAATTTATGCCGTCGATTTTATTTGAAGCGCTGTGCCATATCA 1663  
QY 1757 AAGAAACATTAACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCAACTTTGCTCTTCCTTCC 1816  
DB 1664 AAGCCGAGTTGCTTCAGCGCTGATTTCTGGCGGTGTCTCCAAAGTTCTCTCTCTTCC 1723  
QY 1817 GAGGAATGGAAGCCATTTGAGAGCAATGCAATGGGTTTTCCTTTACCATGCAATCAAGT 1876  
DB 1724 GAGGAATGATCCAGTACGTGAAGCCATTCATGCAAGTGTTCCTCTATTTACGCTATTGCGA 1783  
QY 1877 CTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCCCTGCTGTGATGATGATATCCATA 1936  
DB 1784 ATGGCATGGATATGGGGATTTGTTTACCGCCGCTCAATTTGGCGAATTTATGACACCTGTCTG 1843  
QY 1937 AGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGA 1996  
DB 1844 ATGAATTTAGCTGATGCGGTTCGAGGATGTGATCTCAACCGCGGTGATGACAGTACTGAGC 1903  
QY 1997 AGCTCTTACGTTATGCCAGAC-----TCAAGGCACAGGAGGAAGAAATGTCATTGAGA 2050  
DB 1904 GCCTGCTGATTTAGCAGAAAAATACCGTGCAGATTAAGTAAAGTGGGAGGTCGCGATCCAGC 1963  
QY 2051 CTGATGATGGAGAAATGGCCCTGTGGAAGACCGCTTGATGATGCCCTTGTGAGGGCA 2110  
DB 1964 AAGCAGAAATGGCGTGGTGGCGGTGGTGAACCGTTAGAGTATTCGCTGTGTAAGAGCA 2023  
QY 2111 TTGAAAAATATTAATTCAGGATATCTGAGGAGCCAGGTTAAACCAAAAAATATATCCCC 2170  
DB 2024 TTACTGAAATTTATGAGCTGATACCGAAGAGCGCGCAGCAGGCA-----TC 2074  
QY 2171 GACCTCTCAATATAATTGAGAGACCCCTGATGAATGGAATGGAATTTGTTGGTATCTTT 2230  
DB 2075 GGCCTAATTTGAAGTGAATGAAGGGCGGTTAATGCTCTGGGATGAATGTGTTGTTGCTGCTGT 2134  
QY 2231 TTGGAGCTGGAAGAAATGTTTCTACCTCAGGTTATTAAGTCAAGCTGCGCGGTATGAAGAGG 2290  
DB 2135 TTGTTGAAGGAAATATGTTCTGTCACAGGTTGTAATAATCAGCAGGGTAAATGAAGAGG 2194  
QY 2291 CTGTTGGCCACTTATCCCTTTTCATGGAAAAAGAAAGAGAAACACAGAGTGTCTTAACG 2350  
DB 2195 CGTGGCCCTATCTGAACTTATTTGAA----- 2223  
QY 2351 GCACAGTGAAGAGAGGACCTTACAGGGCACCATCGTGTGGCCACTGTTAAAGCG 2410  
DB 2224 GCCAGTAAGCAGAGGGTACCAACCGCC--GGTAAATCTTATTAGCAACCGCTCAAGGCG 2281  
QY 2411 AGCTGCACATAGGCAAGAACATAGTTGGAGTGTAGTCTTGGCTGCAATTAATTTCCGAG 2470  
DB 2282 ATGTGCATGACATCGGCAAGAACATCGTGGGTGTGGTTTTCAGTGTAAATTAATATGAAA 2341  
QY 2471 TTATTGATTTAGAGTCACTGACTCCATGTGAATAAGTACTGAAAGCTGCTTTGACCAACA 2530  
DB 2342 TCATCGATCTGGCGGTGATGGTCCGACGGAAGAAATTTCTGGCGACTGCAAGGGAAGAGA 2401  
QY 2531 AAGCAGATATAATTTGGCTGTGAGGACTCATCACTCTTCTCCCTGGATGAAATGATTTTGT 2590  
DB 2402 AGGTGGATATATTCGGCTGTCTGGTCTGATTTACCCCGTGTGAGCGAGATGGTTAAGC 2461

QY 2591 TTCCAGGAATATGGAGATTTAGCTATAAGGATTTCCATTTGTTGATTTGGAGGACCAACCA 2650  
DB 2462 TGGGAAAGAGATGGAGCGCCAGGCTTTTACGCTGCGTGTGCTCATTTGGTGGGCAACGA 2521  
QY 2651 CTTCAAAAAACCCACACAGCAGTTTAAATATAGCTCCGAGATACAGTGCACCTGTATTCATG 2710  
DB 2522 CATCCAAAGCTCACACCGCGTCAAGATTGACAAAATTTATAGCGGTTCACCAACGATG 2581  
QY 2711 TCCTGGACGCTCCAAAGATGTGTTGTTGCCAGCTGTTAGATGAATAATCTAAAGG 2770  
DB 2582 TCTCTAACGCTTCGCGAGTGTGGCGTGTGATCCGCTGTGCTATCGGATACGCAACGG 2641  
QY 2771 ATGAATACTTTGAGGAATCATGGAAGATATAGAGATATTTAGACAGGACCATTTATGAGT 2830  
DB 2642 AGGCAATTTGTGGCAAAAAACCGCAAGAAATATGAGACTGTACGTATCCAAC-----ACG 2695  
QY 2831 CTCTCAAGAGAGAGATATCTTACCTTTAAGTCAAGCCAGAAAGTGGTTTCCAATG- 2889  
DB 2696 CCGTTAAAAAACCGCGCACCCCGCGTCAAGTTTTCAGGCTGCACGAAATAACCCCAACGG 2755  
QY 2890 -GATTGCTCTCTGAACCTTCAACAGTGAAGCCACGTTTATTTGGGACCCAGGCTTTGA 2948  
DB 2756 TGATTTGATTTGGAGAACTACACTCCGCACTGGGCCATAAATTTGGGGGTACAGGTGGTGG 2815  
QY 2949 AGACTATGACCTGTCAGAAAGCTGTGGAATACATTTGACTGGAAGCCTTTCTTTGATGCTG 3008  
DB 2816 AGGCCAGTA--TTGAACCTCTGCGCACTATATCGACTGCGCGCATTTCTTTATGACCTG 2873  
QY 3009 GCACTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTAAACGACAAAAACAGT 3068  
DB 2874 GTCGTTGGCAGGCAAGTATCCG-----CGCATTTTGAAGATGAGGTGGT 2918  
QY 3069 AGGTGGAGAGCCAGGAGGCTTACGATGATGCCCAATATGCTGAACACACTGATTAG 3128  
DB 2919 GGGGAGAGGCCCAAGCGCTGCTGCCGACGCAATGCTTTGTTGATAAATTTGTCGCG 2978  
QY 3129 TCAAAAGAAATCTCGGCGCCGGGTGTGGGTCTCTGCCACACAGAGATATCCAAAGA 3188  
DB 2979 CGAGGATCTGTGATCCAAAGGCGTGTGGTCTGTTCCGGCTAACAGCGTCCGCGA 3038  
QY 3189 CGACATTCACCTGTACCCAGAGGCTGTGTGCCCCAGGCTGACAGGCCCATAGCACCTTT 3248  
DB 3039 TGATATCGAAATTTACCGTGTATG-----AGCGGCGAGATGAGGTACTGCGCCATCAG 3089  
QY 3249 CTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGGACCACTACTACTGCT 3308  
DB 3090 CTATCATTTTACGCCAACAGACCCGAAAA-----AACCGACTTCCGAACTACTGTTT 3140  
QY 3309 CTGAGACTTCATCGCTCCCTTGCAATCTGCGCATCCGTGACTTACCTGGGCTGTTTGGCGT 3368  
DB 3141 AGCTGACTATGTCGCGCTTAAATCCAGTGTGTAAGCCGATTTATTTGGTGGTTCGCGT 3200  
QY 3369 TGCC---TGCTTTGGGTAGAAAGAGCTGAGCAAGCCCTATGAGGATGATGTCAGACTA 3425  
DB 3201 GACCGGTGGCTTGGAAAGAGACGATTTGGCCGATGCTTATGATGCGCAACATGATGACTA 3260  
QY 3426 CAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGCGAGAGGCTTTGCGAGAGACT 3485  
DB 3261 CAATAAAATCATGATTAAGGCAATTTATCAGACCGCTGTGCGAAGCCCTTTGCGGAGTATCT 3320  
QY 3486 CCATGAAGAGCTTCGCGAGAACTGTGGGCTCTCTGTGGCAGTGAAGAGCTGGAAGCTGCGC 3545  
DB 3321 GCATGAGGAGTACGTAAGATGATTTGGGTTTGGACCGAAGCAAAATCTGAGCAACGA 3380  
QY 3546 AGACTCGGAAGGTTGGGGTACAAAGGCGATCCGCCCGCTCTCTGCTACCCAGCCAGCC 3605  
DB 3381 AGAGTTGGTGGCGCAAAATTTATCAGGGATTCGCGCCGCGCATCCGGCTATCCGCTTGGCC 3440  
QY 3606 CGACACACCCAGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGAT 3665  
DB 3441 GGAGCAGACCCAGAGAGGCCAAATTTGGCAGTGTGTGGATGTGCAAAACCCACACCGGTAT 3500  
QY 3666 TAGGTTAACAGAAATATTAGCAATAGGCAACCTGCTTACGAGTCTCAGGCTCTACTTCTC 3725

Db 3501 GAAACTGACCGAATCCTACGCCATGTGGCCTTGGGCTTCGGTATCGGTTGGTATTTTTCAG 3560  
Qy 3726 CAATTTGAAGTCCAAATATTTTGTGCTGTGGGGAAGATTTCCAAGGATCAGGTTGAGGATTA 3785  
Db 3561 CCACCCAGACAGCAATATTTTGGCGTAGCACAAATTCAGCGTGATCAGGTTGAGGATTA 3620  
Qy 3786 TGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTTTGGG 3845  
Db 3621 TGCCGCCCGTAAAGGGATGCCCACTGCCGAAGTTGAACGCTGGCTAGCCCTAATCTGGG 3680  
Qy 3846 ATATGATACAGACTAA 3861  
Db 3681 TTACGATCGGATTTGA 3696

Search completed: March 6, 2006, 19:35:54  
Job time : 2013 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:02:53 ; Search time 698 Seconds  
(without alignments)  
12309.956 Million cell updates/sec

Title: US-10-607-712-1  
Perfect score: 3919  
Sequence: 1 ggcacactgaggagacg.....ctcaaggaaatacaactag 3919

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:

- 1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:
- 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:
- 5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:
- 7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:
- 9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:
- 10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:
- 11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:
- 12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883.8	22.6	3684	US-10-858-730-261	Sequence 261, App
C 2	141.4	3.6	1275	US-09-925-065A-727226	Sequence 727226, App
C 3	135.8	3.5	494	US-11-116-881A-1611	Sequence 1611, Ap
4	112.8	2.9	2599	US-11-055-822-211	Sequence 211, App
5	112.8	2.9	2599	US-11-055-822-709	Sequence 709, App
6	112.8	2.9	3666	US-10-858-730-260	Sequence 260, App
7	112.4	2.9	2578	US-11-055-822-213	Sequence 213, App
8	112.4	2.9	2578	US-11-055-822-711	Sequence 711, App
C 9	101.4	2.6	616	US-09-925-065A-697175	Sequence 697175, App
10	93.4	2.4	3477	US-10-858-730-141	Sequence 141, App
11	91.4	2.3	3513	US-10-858-730-142	Sequence 142, App
12	89.6	2.3	3621	US-10-858-730-144	Sequence 144, App
13	88.2	2.3	3579	US-10-858-730-143	Sequence 143, App
14	46.4	1.2	50	US-11-175-859-4373	Sequence 4373, Ap
C 15	43.2	1.1	616	US-09-925-065A-483679	Sequence 483679, App
C 16	41	1.0	588	US-09-925-065A-39453	Sequence 39453, A
C 17	40.4	1.0	1877	US-09-925-065A-671655	Sequence 671655, App
C 18	40	1.0	119036	US-10-995-561-13314	Sequence 13314, A
C 19	39.6	1.0	591	US-09-925-065A-408435	Sequence 408435, App
C 20	39.6	1.0	731	US-09-925-065A-680394	Sequence 680394, App

21	39.6	1.0	731	6	US-09-925-065A-680395	Sequence 680395, App
22	39.6	1.0	160226	12	US-11-121-086-29	Sequence 29, Appl
C 23	39	1.0	611	6	US-09-925-065A-460077	Sequence 460077, App
C 24	39	1.0	1284	12	US-11-098-686-9926	Sequence 9926, Ap
25	39	1.0	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
C 26	38.8	1.0	354	6	US-09-925-065A-125604	Sequence 125604, App
C 27	38.8	1.0	1100	8	US-10-750-185-25373	Sequence 25373, A
C 28	38.8	1.0	1100	8	US-10-750-623-25373	Sequence 25373, A
29	38.2	1.0	179487	7	US-10-330-773-664	Sequence 664, App
30	37.6	1.0	591	6	US-09-925-065A-281175	Sequence 281175, App
31	37.6	1.0	21442	8	US-10-995-561-13469	Sequence 13469, A
C 32	37	0.9	499	6	US-09-925-065A-594374	Sequence 594374, App
C 33	37	0.9	580	6	US-09-925-065A-215989	Sequence 215989, App
C 34	37	0.9	591	6	US-09-925-065A-816505	Sequence 816505, App
C 35	37	0.9	319608	12	US-11-145-703-1	Sequence 1, Appl
36	36.8	0.9	2475	12	US-11-136-527-3385	Sequence 3385, App
37	36.6	0.9	200	12	US-11-098-686-7768	Sequence 7768, App
C 38	36.6	0.9	499	6	US-09-925-065A-594373	Sequence 594373, App
C 39	36.6	0.9	588	6	US-09-925-065A-428604	Sequence 428604, App
C 40	36.6	0.9	571	6	US-09-925-065A-330747	Sequence 330747, App
41	36.6	0.9	573	6	US-09-925-065A-598361	Sequence 598361, App
42	36.6	0.9	602	6	US-09-925-065A-384546	Sequence 384546, App
C 43	36.6	0.9	1273	8	US-10-750-185-41055	Sequence 41055, A
C 44	36.6	0.9	1273	8	US-10-750-623-41055	Sequence 41055, A
45	36.6	0.9	2141	6	US-09-925-065A-685376	Sequence 685376, App

ALIGNMENTS

RESULT 1

US-10-858-730-261  
; Sequence 261, Application US/10858730  
; Publication No. US20050255568A1  
; GENERAL INFORMATION:  
; APPLICANT: Bailey, Richard B.  
; APPLICANT: Blomquist, Paul  
; APPLICANT: Doten, Reed  
; APPLICANT: Driggers, Edward M.  
; APPLICANT: Madden, Kevin T.  
; APPLICANT: O'Leary, Jessica  
; APPLICANT: O'Toole, George  
; APPLICANT: Trueheart, Joshua  
; APPLICANT: Walbridge, Michael J.  
; APPLICANT: Yorgey, Peter S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID  
; TITLE OF INVENTION: PRODUCTION  
; FILE REFERENCE: 14184-030001  
; CURRENT APPLICATION NUMBER: US/10/858,730  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR FILING DATE: 2003-05-30  
; PRIOR FILING DATE: 2004-03-10  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 261  
; LENGTH: 3684  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-10-858-730-261

Query Match	22.6%	Score 883.8	DB 8	Length 3684
Best Local Similarity	54.4%	Pred. No. 2.2e-238		
Matches 2043	Conservative	0	Mismatches 1612	Indels 102
Gaps				9
QY	120	GGATGAGATCAATGCCATTCTGCAGAGAGAGATTATGGTCTGGATGGAGGATGGGAC	179	
Db	15	GGACACTCGTGGCGGATTAATGAACGTTATCTGGTCTGACGCGGATGGGAC	74	
QY	180	CATGATCCAGCGGAGAGAGCTAAACGAAGAACCTCCAGAGGTGACAGAAATTAAGATCA	239	
Db	75	CATGATCCAGAGTATTCGACTGAAAGCGGATTTTCGTGGTGAACGCTTTGCCGACTG	134	



```
Db 2247 GGTGATCGCCACCGTGAAGCGCGACGCTCCACGACATCGGTAAATAATATCGTGGTGGT 2306
Qy 2448 CCTTGGCTGCAATAATTTCCGAGTTATTTAGTATTTAGAGTCAATGACTCCATGTGATAAGAT 2507
Db 2307 GCTGCAATGTAAACAATACGAAATTTGCGATCTCGGGTTATGGTCTCGCGGAAATAAT 2366
Qy 2508 ACTGAAGAGCTCTTTCACACAAAGCAGATATAATTTGGCTGTGAGGACTCATCACTCC 2567
Db 2367 TCTCCGTAACCGCTAAAGAAAGTGAATGCTGATCTGATTTGGGCTTTTCGGGGCTTATCACGCC 2426
Qy 2568 TTCCTCGATGAATGATTTTGTGCAAGGAAATGAGAGATTTAGCTATAAGGATTC 2627
Db 2427 GTGCTGAGCAGATGTTTAACTGCGGAAAGAGATGAGCGGTGAGGGCTTCAATATCC 2486
Qy 2628 ATTGTTGATTTGAGGAGCAACCACTTCAAAAACCCACACAGCAGTTTAAATAAGCTCCGAG 2687
Db 2487 GTTACTGATTTGGCGCGCAGCACCTCAAAGCGCACACCGCGGTGAATAATCGAGCAGAA 2546
Qy 2688 ATACAGTGCACTGTAATCATCTGCTGGAGCGGTCCAGAGTGTGGTGTGTTGCCCA 2747
Db 2547 CTACAGCGCGCCGACGGTGTATGTCAGAAATGCTCTGCGTACCGTTGGTGTGGCGGC 2606
Qy 2748 GCTGTTAGATGAATACTAAAGATGAATCTTTGAGGAATCATGGAAGAAATATGAAGA 2807
Db 2607 GCTGCTTTCCGATACCCAGCGTGATGTTTGTGCTGTGACCGCGAAGGAGTACGAAAC 2666
Qy 2808 TATTAGACAGGACCATTTATGATCTCTCAAGAGAGAGAGATACTTACCCCTTAAGTCAAGC 2867
Db 2667 CGTACGTATTTCAGCACGGCGCGCAAGAAACCGCGCACACACCGGTCAAGTGAAGCGGC 2726
Qy 2868 CAGAAAAGTGGTTTCCAAATGAGTGGCTGTCTGAACCTCACCAAGTGAAGCCCAAGTT 2927
Db 2727 GCGCGATAACGATTTGCTTTTGAAGTGGAGGCTTACACGCCCGGTGGCGCACCGTCT 2786
Qy 2928 TATTGGACCCAGTCTTTGAAGACTATGACCTGCGAGAACTGCTGAGACTGACTGACTG 2987
Db 2787 CGCGGTGACGAAATC-----GAAGCCAGCATCGAAACCGTCTGTAATTTACATCGACTG 2840
Qy 2988 GAAGCCCTTCTTTGATCTGTCGAGCTCGCGGGCAAGTACCCGAACTGAGGCTTCCCCAA 3047
Db 2841 GACACCGTCTTTATGACCTGTCGTCGTCGCGGGAGATATCG-----CG 2885
Qy 3048 GATATTTAACACAAAACAGTAGTGGAGGCGCAGGAAGGTCTACGATGATGCCACAA 3107
Db 2886 CATCTGGAAGATGAAGTGGGCGTTGAGGCGCAGCGCTGTTTAAAGACGCCAACGA 2945
Qy 3108 TATGCTGAACACACTGATTAGTCAAAAGAACTCCGCGCCGCGGTGGTGGTCTG 3167
Db 2946 CATGCTGGATAAATTAAAGCCCGGAGAAACCGCTGAATTCGCGGTGGTGGGCGCTGTT 3005
Qy 3168 GCCAGCACAGATATCCAGACGACATTTACCTGTACGAGAGGCTGCTGTGCCCGCAGGC 3227
Db 3006 CCGCGCAACCGTGTGGCGATGACATTAATCTACCGTGACGAAACGCGTACCCATG- 3064
Qy 3228 TGCAGAGCCCATAGCCACTTTCTATGGGTTTAAAGCAACAGGCTGAGAGGACTCTGCCAG 3287
Db 3065 -----TGATCAAGCTCAGCCACCATCTCGTCAACAGACCGGAAATAAC 3107
Qy 3288 CACGAGCCATATCTATGCTCTCAGACTTCTGCTCCCTTGTGATTTCTGGCATTCGCTGA 3347
Db 3108 AGGCTTCGCTAACTACTGCTCGTGAATCTGTTGCGCCGAAAGCTTTCTGTTAAAGCAG 3167
Qy 3348 CTACTCGGCTGTTTCCGCTGCTGCTTTTGGGGTAGAAGAG---CTGAGCAGGCGCTA 3404
Db 3168 TTACATCGGCGCATTTTCCCGTGACTGCGGGGCTGGAAGAGGACGCACTGGCTGATGCCIT 3227
Qy 3405 TGAGGATGATGGTGACGACTACAGCAGCATCATGCTCAAGCGCTGAGGAGCCGCGTGGC 3464
Db 3228 TGAAGCCGACGACGATGATTACACAAATATGTTGAAGAGCGTTGCCGACCGTTTACG 3287
Qy 3465 AGAGGCTTTTGCAGAAAGAGCTCCATGAAGAGTTCGCGCAGAACTGTGGGCGCTACTGTGG 3524
```

```
Db 3288 CGAAGCCTTTGGAGTATCTCCATGAGCGGTGTGCGTAAAGTCTACTGGGGCTATGCGCC 3347
Qy 3525 CAGTGAAGAGCTGGAGCTGCGAGACCTGCGAAGTTTGCAGTACAAAGGCGATCCGCCCGGC 3584
Db 3348 GAACGAGAACCTCAGCAACGAAAGAGCTGATCGCGAATAACTACGAGGCGATCGTCCGGC 3407
Qy 3585 TCCTGCTACCCAGCAGCCGACACACCGAGAGCTCACCATGTGGAGACTCGCAGA 3644
Db 3408 ACCGGCTATCGGCGCTGCGCGGAAACATACGGAATAAGCCACATCTGGAGCTGTGGA 3467
Qy 3645 CATCGAGCAGTCTACAGGCAATTAGGTTAAACAGAAATCATTTAGCAATGGCACCTGCTTCAGC 3704
Db 3468 AGTGAAGAAACACACTGGCATGAACTCACAGAACTTTTCGCCATGTGGCCCGGTGATC 3527
Qy 3705 AGTCTCAGCGCTCTACTCTTCCAAATTTGAAGTCCAAATAATTTTCTGTGGGGAAGATTTC 3764
Db 3528 GGTTCGGGTGTGTTACTTTCAGCCACCGGACAGAGTACTACGCTGTAGCACAAATTCA 3587
Qy 3765 CAAGGATCAGTTGAGGATTTATGATTTGGGATATGATACAGACTAA 3861
Db 3588 GCGCGATCAGTTGAAGATTATGCGCGCGTAAAGGTATGAGCGTTACCGAAGTTGAGCG 3647
Qy 3825 ATGGCTTGACCCATTTTGGGATATGATACAGACTAA 3861
Db 3648 CTGGCTGGCACCGATCTGGGGTATGACCGGACTCA 3684
```

## RESULT 2

US-09-925-065A-727226/c

Sequence 727226, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 727226

LENGTH: 1275

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-727226

Query Match 3.6%; Score 141.4; DB 6; Length 1275;

Best Local Similarity 99.3%; Pred. No. 6.1e-29;

Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## RESULT 3

US-11-116-881A-1611/c  
; Sequence 1611, Application US/111116881A  
; Publication No. US20060041949A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Dongmei  
; APPLICANT: Nielsen, Mark T.  
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof  
; FILE REFERENCE: 07678/141014  
; CURRENT APPLICATION NUMBER: US/11/116,881A  
; CURRENT FILING DATE: 2005-04-27  
; PRIOR APPLICATION NUMBER: 60/665,451  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/665,097  
; PRIOR FILING DATE: 2005-03-24  
; PRIOR APPLICATION NUMBER: 60/646,764  
; PRIOR FILING DATE: 2005-01-25  
; PRIOR APPLICATION NUMBER: 60/607,357  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 60/566,235  
; PRIOR FILING DATE: 2004-04-29  
; PRIOR APPLICATION NUMBER: 10/934,944  
; PRIOR FILING DATE: 2004-09-03  
; PRIOR APPLICATION NUMBER: 10/943,507  
; PRIOR FILING DATE: 2004-09-17  
; PRIOR APPLICATION NUMBER: 60/503,989  
; PRIOR FILING DATE: 2003-09-18  
; PRIOR APPLICATION NUMBER: 60/485,368  
; PRIOR FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: 60/418,933  
; PRIOR FILING DATE: 2002-10-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2300  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1611.  
; LENGTH: 494  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-11-116-881A-1611

Query Match 3.5%; Score 135.8; DB 9; Length 494;  
Best Local Similarity 56.2%; Pred. No. 1.3e-27;  
Matches 276; Conservative 0; Mismatches 212; Indels 3; Gaps 1;  
  
QY 260 ACAATGACATTTTAAGTAACTAGCTGATGATTCATTAACCAATCCATTAAGCAATCT 319  
DB 494 ACAAGACCTGCTGGTACTCAGTAAACCGGAAGTATCGCGCTATCCACACGCTACT 435  
  
QY 320 TGCTGGCTGGGCGAGATATCATTGAAACAAATACCTTTAGCAGCACTAGTATGCCCAAG 379  
DB 434 TTGAAGCGGCGCGGATATCATCGAAACCAACACCTTCACTCCACGACCATTTGGATGG 375  
  
QY 380 CTGACTATGGCTTGAACACATTTGGCTTACCGGATGAACATGTGCTCTGCAGAGTGGCCA 439  
DB 374 CGGATTACAGATGAATCCCTGTGCGCGGAAATCACTTTGCGCGCGGAAACTGGCGC 315  
  
QY 440 GA---AAGTGTCCGAGAGGTAACTCTCCAGACAGGAATTAAGAGTTTGTGCGAGGG 496  
DB 314 GAGCTTGTGCTGACAGTGGGACCGCGCGCACCGCAGAGAGAAACCGCGCTACGTTGCCGGTG 255  
  
QY 497 CTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCATCTGTGAAAGCGCGGATTATA 556  
DB 254 TTCTCGGCGCGCAACCGCAGCGGCTCTATTCTCCGACGTCAACGATCCGCAATTC 195  
  
QY 557 GGAAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCCAAAGACTTTCTGG 616  
DB 194 GTAAATACATTTTGACGGCTGTGGCGGCTTATCGAGAGTCCACCAAGGCGCTGGTG 135  
  
QY 617 ATGGCGGGTGTATCTTACTCATTTGAACACTATTTTGTATCTGCCATGCCAAGCAG 676  
DB 134 AAGTGGCGCGGATCTGATCTCTGATGAACACCGTTTTCGACACCTTAAACGCAAGCGG 75  
  
QY 677 CCTTGTGTGCACTCCAAAATCTTTTGTAGGAGAAATATGCTCCCGGCGCTATCTTTATT 736

DB 74 CGGTATTTGCGGTGAATAACCGAGTTTGAAGCGCTGGCGGTTGAGCTGCCGATTATGATCT 15  
QY 737 CAGGAGCATC 747  
DB 14 CCGGCACCATC 4

## RESULT 4

US-11-055-822-211  
; Sequence 211, Application US/11055822  
; Publication No. US20050260707A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zeider, Oskar  
; APPLICANT: Haberkauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
; FILE REFERENCE: BGI-121CPCN  
; CURRENT APPLICATION NUMBER: US/11/055,822  
; CURRENT FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: 09/606,740  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/141,031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 60/142,101  
; PRIOR FILING DATE: 1999-07-02  
; PRIOR APPLICATION NUMBER: 60/148,613  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: 60/187,970  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: DE 19930476.9  
; PRIOR FILING DATE: 1999-07-01  
; PRIOR APPLICATION NUMBER: DE 19931415.2  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931418.7  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931419.5  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931420.9  
; PRIOR FILING DATE: 1999-07-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1158  
; SEQ ID NO 211  
; LENGTH: 2599  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)..(2599)  
; OTHER INFORMATION: RXN02198  
US-11-055-822-211

Query Match 2.9%; Score 112.8; DB 12; Length 2599;  
Best Local Similarity 45.4%; Pred. No. 1.2e-20;  
Matches 650; Conservative 0; Mismatches 737; Indels 45; Gaps 5;  
  
QY 1162 CCGTACCAACCTTTTGTAACTTTGGAGCGCTGTAAATGTTGCGAGTCAAGGAAGTTT 1221  
DB 1187 CAGGAACCGGCAATTTCCATGATCGGTGAGCGCAACCACTCCACGGTTCCAGGCATTC 1246  
  
QY 1222 GCTAACTCATCATGCGAGGAACATATGAAGACCTTGTGTGTCCTCAAGTGCAGGTG 1281  
DB 1247 CGTGAGCAATGCTGTCTGCGGATTTGGGAAAGTGTGTGATATTGCCAAGCAGCAACC 1306  
  
QY 1282 GAAATGGAGCCCGAGGTGTTGGATGTCAACATGATGATGGCATGTGTAGATGGTCCAAAGT 1341  
DB 1307 CGCATGTGACACATGCTGGATCTTTGTGTGATTTACGTGGGACGAGCGCACCGCC 1366  
  
QY 1342 GCAATGACCAAGATTTTGCACACTTAATTTGCTTCCGAGCCAGACATCGCAAGAGGTACTTTG 1401

1367 GATATGGCGACCTTTGGCAGCACTTCTTGCTAC-----CAGCTCCACTTTTGCCAAATC 1417  
1402 TGCATCGACTCTCTCAATTTTCTGTGATTTGAAGCTGGTTAAAGTGTGCTGCCAAGGAAG 1461  
1418 ATGATTGACTCCACCGGACGAGGTTATTGCGACAGCGCTTGAGCACTTTGGGTGGACGA 1477  
1462 TGCATTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCA-----CTTCTTGGAG 1512  
1478 AGCATGTCTTACTCCGTCACTTTGAAGAGCGCGATGGCCCTGAGTCCCGCTACACGCG 1537  
1513 AAGGCCAGGAAGATTAAAAAGTATGGAGCTCTATTTGGTGTGATGGCTTTTGTATGAAGAA 1572  
1538 ATCATGAATCTGTAAAGCAGCAAGTGGCGCTGTGGCTGACCAATGATGAGGAA 1597  
1573 GGACAGGCAACAGAAACAGACAAATAATAGAGTGTGACCGCGGCTTACCATCTGCTT 1632  
1598 GGCACGGCAGCATGCTGCTGAGCACAAGGTGGCAATGCTAAACGACTGATTCAGCAATC 1657  
1633 GTGAAAAAATCGGCTTTAATCCAAATGACATTAATTTTGAACCCCTAATATCCTAACCAT 1692  
1658 ACCGCGAGCTAGCGCTGGATATCAAGACATCTGTTGGACTGCTGACCTTCCCGATC 1717  
1693 GGGATGGAATGGAGGAACAACATTTGATGCTCAATTAATTTTCCATGCAACAAAGATC 1752  
1718 TCTACTGCCAGGAAGAAACAGCGGAGATGGCAATGAACCATGGAAGCATCCGCGAG 1777  
1753 ATTAAGAAACATTAATCTGGAGCAGAAATAGTGGAGTCTTTCACATCTGCTCTCC 1812  
1778 CTGAAGAAGCTCTACCCAGAAATCCACACACCTGGTGTGCTCAATATTTCTTCCGCG 1837  
1813 TTCCGGAATGGAAGCAATTCGAGAGCAATGATGGGTTTTCCTTTACCTGCAATC 1872  
1838 CTG-----AACCTGTGCGACCGCAGGTTCTTAATCTGTGTTCTCAATGATGCAAT 1891  
1873 AAGTCTGGCATGAGATGAGTGAATGCTGAAACCTCCCTGTGTATGATGATC 1932  
1892 GAGGCTGTCTGACTCTGCAATTCGCAAGCTCCAGATTTGCGATGACCGCAT 1951  
1933 CATGAAGAACTTCTGAGCTCTGTAAGATCTCATCTGGAATAAAGACCTTGAGGCGCACT 1992  
1952 GATGATCCAGCGGAGTGGCGTTGATGCTGATGATGCTGATGCTGCGCAGCGAGATTAC 2011  
1993 GAGAGCTCTTAATGTTATGCCAGCTCAAGGACAGGAGGAAGATCAATCAGACT 2052  
2012 GATCGCTGCGAGAAATCATGCACTCTTTGA-----GGCGGTTTCTGCTGCGGATGCCAAG 2068  
2053 GATGAGTGGAGAAATGGCCCTGTGGAAGACGCTCTGATGATGCTGTTGGAAGGCAAT 2112  
2069 GATGCTGCGCTGAACAGCTGCGCGCTATGCTTTGTTGAGCGTTTGGCAGCGCATC 2128  
2113 GAAAAACATATTAATGAGGATCTGAGGAAGCCAGGTTAAACCAAAAAAATAATCCCGA 2172  
2129 ATCGACGCGGATAAGATGGCTTTGAGGATGATCTGGAAGCAGCATGAAGGAAGTCT 2188  
2173 CCTCTCAATATAATGGAAGACCCCTGATGATGGAATGGAATTTGTTGATCTTTT 2232  
2189 CCTATTGCGATCATCAACAGGAGCTTTCTCAACGCGCATGAAGACCGTGGGTGAGCTGTT 2248  
2233 GGAGCTGGAATAATGTTCTTACCTCAGCTTATAAGTCAAGCCCGGTTATGAAGAGGCT 2292  
2249 GGTTCGGAACAGATGAGCTGCCATTTCTGCTGCAATCGGAGAAACCAATGAATCTGCG 2308  
2293 GTTGGCCACCTTATCCCTTTTCATGAAAAAAGAGAGAGAAACAGAGTGTCTTAACGCG 2352  
2309 GTGGCTTATTGGAACCGTTTCAATGAAGAGGAAGCAGAGCTACCGGATCT----- 2359  
2353 ACAGTAGAAGAGAGACCTTTACAGGCAACCATCTGCTGGCCATCTGTTAAAGCGAC 2412  
2360 -----GGCAGGAGGAGGGAAGGCAAAATCGTGTGGCCACCGTCAAGGGTGAC 2410  
2413 GTGCACGACATAGCAGACATAGTTGAGTAGTCTTCTGCTGCAATATTTCCGAGTT 2472  
2411 GTGCACGATATCGGCAAGAACTTGTGGACATCATTTTGTCCAAACACCGTTACGACGTG 2470

QY 2473 ATTGATTAGGACTCATGACTCATGTGATAGATGACTGAAAGCTGCTTCTTGACCACAA 2532  
Db 2471 GTGAACCTTGGGCATCAAGCAGCCACTGTCCGCCATGTTGGAAGCAGCGGAAGACACAA 2530  
QY 2533 GCAGATATTAATTCGCTGTCTCAGGACTCATCACTCTCTTCCCTGATGAAATGA 2584  
Db 2531 GCAGAGCTCATCGGCATGTCGGGACTTCTTGTGAAGTCCACCGTGGTGATGA 2582

## RESULT 5

US-11-055-822-709  
; Sequence 709, Application US/11055822  
; Publication No. US20050260707A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Habershauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
; FILE REFERENCE: BGI-121CPCN  
; CURRENT APPLICATION NUMBER: US/11/055,822  
; CURRENT FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: 09/606,740  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/141,031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 60/142,101  
; PRIOR FILING DATE: 1999-07-02  
; PRIOR APPLICATION NUMBER: 60/148,613  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: 60/187,970  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: DE 19930476.9  
; PRIOR FILING DATE: 1999-07-01  
; PRIOR APPLICATION NUMBER: DE 19931415.2  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931418.7  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931419.5  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931420.9  
; PRIOR FILING DATE: 1999-07-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1158  
; SEQ ID NO 709  
; LENGTH: 2599  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)..(2599)  
; OTHER INFORMATION: RXN02198  
US-11-055-822-709

Query Match 2.9%; Score 112.8; DB 12; Length 2599;  
Best Local Similarity 45.4%; Pred. No. 1.2e-20;  
Matches 650; Conservative 0; Mismatches 737; Indels 45; Gaps 5;  
QY 1162 CCGTACACCAACTTTGTTAACTTGGAGAGCGCTGTAATGTTGCGAGATCAAGGAAGTTT 1221  
Db 1187 CAGGAAACCGGCATTTCCATGATCGGTGAGCGCACCAACTCCACCGTTTCCAAGGCATTC 1246  
QY 1222 GCTAAACTCATATGCGCAGGAAGTATGAGAGCCCTTGTGTTGCCAAGTGCAGGTG 1281  
Db 1247 CGTGAGCAATGCTGTCTGCGGATTCGGGAAAGTGTGTGATATTGCCAAGCAGCAACC 1306  
QY 1282 GAAATGGAGCCCGCAGGTGTTGGATGTCAACATGATGATGGCATGTAGATGTGCAAGT 1341  
Db 1307 CGCGATGTTGACACATGCTGGATCTTTGTGTGGATTTACGTGGGACGAGCGCACCGCC 1366

```
Qy 1342 GCAATGACCGAGATTTTGCAACTTAATTGCTTCCGAGCCAGACATCGAAAGGTACCTTTG 1401
Db 1367 GATATGGGACCTTGGCAGCACTTCTTGCTAC-----CAGCTCCATTTGCCAATC 1417
Qy 1402 TGCATCGACTCTCCAAATTTTGTGTGATGAAGCTGGTTAAAGTGTCTGCAAGGGAAG 1461
Db 1418 ATGATTGACTCCACCGAGCCAGAGGTTATTGCGCACAGGCCCTTGAGCACTTGGGTGGACGA 1477
Qy 1462 TGCATTTGTCATAGCATTTAGTCTGAAGGAGGAGGACGA-----CTTCTTTGGAG 1512
Db 1478 AGCATCGTTAACTCCGTCGAACCTTTGAAGAGCGGATGGCCCTGAGTCCCGCTACACGCGC 1537
Qy 1513 AAGCCAGGAGAGATTTAAAGATATGGAGCTGCTATGGTGGTCATGGCTTTTGTATGAAGAA 1572
Db 1538 ATCATGAACCTGGTTAAGCAGACAGGTGGCGCCGTGGTTGGCTGACCAATTTGATGAGGAA 1597
Qy 1573 GGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGSCACCCGGGGCTACCAATCTGCTT 1632
Db 1598 GGCCAGGCACGTACCGCTGAGCACAAAGGTGGCATTTGCTAAACGACTGATTGACGATATC 1657
Qy 1633 GTGAATAAACTGGGCTTTAATCCAAATGACATTAATTTTGACCCCTAATATCTTAACCAATT 1692
Db 1658 ACCGGCAGCTACGGCTGGATATCAAGACATCGTTGTGGACTGCGTGAACCTTCCCGATC 1717
Qy 1693 GGGACTGGAATGGAGGAAACAACTTGTATGCCATTAATTTTATCCATGCAACAAAAAGTC 1752
Db 1718 TCTACTGGCCAGGAGAAACACGCGAGATGGCAATGAAACCATGGAAGCATCCGCGAG 1777
Qy 1753 ATTAAGAAACATTAACCTGGAGCCAGATAAGTGGAGGTCTTTCCAACTGTCTCTTTCTCC 1812
Db 1778 CTGAAGAAGCTCTACCCAGAAATCCACACCACCTGGGTCTGTCCAAATATTCTCTCGGC 1837
Qy 1813 TTCGAGGAATGGAAGCAATTCGAGAACATTCGAGAACATGATGGGGTTTCCITTACCATGCAATC 1872
Db 1838 CTG-----AAACCTGTGTCAGCGCATGTTCTTAATCTGTGTTTCTCAATGAGTGCAAT 1891
Qy 1873 AAGCTGCGATGACATCGAGATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATC 1932
Db 1892 GAGGCTGTCTGATCTGCGATTTGCGCATTCGACAGCTCCAGATTTTGCCGATGAACCGCAT 1951
Qy 1933 CATAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACT 1992
Db 1952 GATGATGCCAGCGGAGTGGCGTTGGATATGGTCTATGATCGCGCACCGAGGATTAC 2011
Qy 1993 GAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGAGGAAAGTCAATTCAGACT 2052
Db 2012 GATCGCTGCGAGGAATTCATGACGCTGTTTGA---GGGCGTTTCTGCTGCCGATGCCAAG 2068
Qy 2053 GATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCTTTGTGAAGGCAAT 2112
Db 2059 GATGCTCGCTGAACAGCTGCGCGCTATGCTTTGTTGAGCGTTTGGCACAGCGCATC 2128
Qy 2113 GAAAAACATATTTATTGAGGATATCTGAGGAAGCCAGGTTTAAACCAAAAAAATATCCCGGA 2172
Db 2129 ATCGAGCGGATAGAATGGCTTGAGGATGATCTGGAAGCAGGCGATGAAGAGAGTCT 2188
Qy 2173 CCTCTCAATATTAATGAAGAGCCCTGATGAATGAATGAATAATTTGTTGGTGAATCTTTT 2232
Db 2189 CCTATTGGATCATCAACAGGAGACCTTCTCAACGCGCATGAAGACCGTGGGTGAGCTGTTT 2248
Qy 2233 GGAGCTGAAAAATGTTTCTACTCAGGTTTATAAAGTCAAGCCGCGGTTTATGAAGAAGCT 2292
Db 2249 GGTTCGGAACAGATGAGCTGCGCATTCGTGCTGCAATCGGCAGAAACCATGAATACTGG 2308
Qy 2293 GTTGCCACCTTATCCCTTTTCATGAAAAAGAGAGAAAGCAAGAAACAGAGTGTCTTAAACGCG 2352
Db 2309 GTGGCTATTTTGGAACCGTTTCATGAAGAGGAGGAGGAGGAGTACCGGATCT----- 2359
Qy 2353 ACAGTAGAAGAGAGACCTTACAGGGCACCATCTGCTGCTGGCCACTGTTTAAAGGCGAC 2412
Db 2360 -----GGCAGGCGAGAGGGCAAGGGCAAAATCGTGGCCACCGTCAAGGGGTGAC 2410
Qy 2413 GTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATAATTTTCCGAGTT 2472
```

```
Db 2411 GTGCAGATATCGGCAAGAACTTTGGTGGACATCATTTTGTCCAAACACGGTTAGCAGGTG 2470
Qy 2473 ATTGATTTAGAGTCATGACTCCATGTGATAGATCTGAAAGTGTCTTGTGACCACAAA 2532
Db 2471 GTGAACCTTGGGCAATCAAGCAGCCACTGTCCGCCATGTTGGAAGCAGCGGAGAAACACAAA 2530
Qy 2533 GCAGATATAATTCGCCCTGTCCAGGACTCATCACTCTCTTCCCTGGATGAAATGA 2584
Db 2531 GCAGACGTATCGGCATGTCCGGACTTCTTGTGAAGTCCACCGTGGTGATGA 2582

RESULT 6
US-10-858-730-260
; Sequence 260, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 3666
; TYPE: DNA
; ORGANISM: Coryne-bacterium glutamicum
US-10-858-730-260

Query Match 2.9%; Score 112.8; DB 8; Length 3666;
Best Local Similarity 45.3%; Pred. No. 1.5e-20;
Matches 659; Conservative 0; Mismatches 752; Indels 45; Gaps 5;

Qy 1162 CCGTACACCAACTTTGTTAAACATTTGGAGAGCGCTGTAAATGTTGAGGATCAAGGAAGTTT 1221
Db 1087 CAGGAAACCGGCATTTTCATGATCGGTGAGGCAACCAACTCCACGTTCCNAGGCATTTC 1146
Qy 1222 GCTAAATCATCATGCGAGGAAACTATGAAGAACCTTGTGTGTTGTTGCAAGTGCAGGTG 1281
Db 1147 CGTGAGGCAATGCTGTCTGGCGATTGGGAAAGTGTGTGGATATTGCGACAGCAAAACC 1206
Qy 1282 GAAATGGAGCCAGGTGTTGGATGTCAATGATGATGCGATGCTAGATGGTCCAGT 1341
Db 1207 CGCATGCGTGCACATGCTGGATCTTTGTGTGATTTACGTGGGACGAGCGCACCGCC 1266
Qy 1342 GCAATGACCAAGATTTTGCMACTTAATTTGCTTCCGAGCCAGACATCGAAAGGTACCTTTTG 1401
Db 1267 GATATGGGACCTTGGCAGCACTTCTTGCTAC-----CAGTCCACTTTGCCAATC 1317
Qy 1402 TGCATCGACTCTCCAAATTTTGTGTGATGAAGCTGGTTAAAGTGTCTGCAAGGGAAG 1461
Db 1318 ATGATTGACTCCACCGAGCCAGAGGTTATTGCGACAGGCCCTTGAGCACTTGGGTGGACGA 1377
Qy 1462 TGCATTTGTCATAGCATTTAGTCTGAAGGAGGAGGAGGACGA-----CTTCTTGGAG 1512
Db 1378 AGCATCGTTAACTCCGTCGAACCTTTTGAAGAGCGGCGATGGCCCTGAGTCCCGCTCAAGCGC 1437
Qy 1513 AAGGCCAGGAAGATTAAAAAGTATGGAGTCTGCTATGGTGGTTCATGGCTTTTGTATGAAGAA 1572
```





```
Qy 1462 TGCATTGTCAATAGCAATAGTCTGAAGGAAGGAGGACGA-----CTTCTTGAG 1512
Db 1478 AGCATGTGTAACCTCGTCAACTTTGAAGACGGGATGGCCCTGAGTCCCGCTACAGGC 1537
Qy 1513 AAGCCAGGAGAGATTAAAGATATGGAGCTCTATGGTGTGTCATGGCTTTTGTATGAAGAA 1572
Db 1538 ATCATGAACTGGTTAAAGCAGACCGTGGCGCCGTTGGTGGCTGACCACTTGTATGAGGAA 1597
Qy 1573 GGACAGGCAACAGAAACAGACACAAAATCAGAGTGTGCACCCGGGGCTACCATCTGCTT 1632
Db 1598 GGCCAGGCAACGATCCGCTGAGCACAAGGTGGCATTTGCTAAACGACTGATTGACGATATC 1657
Qy 1633 GTGAAAAAATCGGCTTTAATCCAAATGACATTTATTTTGACCCCTAAATATCTTAACCAAT 1692
Db 1658 ACCGGCAGCTACGGCTGGATATCAAGACATCTGTTGGACTGCTGACCTTCCGATC 1717
Qy 1693 GGGACTGGAATGGAGGAACACAACTTGTATGCCATTAATTTTATCCATGCAACAAAAGTC 1752
Db 1718 TCTACTGGCCAGGAAGAAACAGCGAGATGGCATTTGAACCATCGAAGCCATCCGCGAG 1777
Qy 1753 ATTAAAGAACTTACCTGGAGCCAGATAAGTGGAGTCTTTTCAACTTGTCTTCTTCC 1812
Db 1778 CTGAAGAAGCTCTACCCAGAAATCCACACCACTGGGTCTGTCCAAATATTTCTCTCGGC 1837
Qy 1813 TTCCGAGGATGGAAGCCATTCGAGAAGCAATGATGGGTTTTCTTTTACCATGCAATC 1872
Db 1838 CTG-----AACCTGTGCGACCGCAGGTTCTTAATCTGTGTTCTCTCAATGATGCAAT 1891
Qy 1873 AAGTCTGGCATGAGATAGTGAATCTCGAAACCTTCCCTGTGTATGATGATATC 1932
Db 1892 GAGGCTGTCTGGAATCTGCGATTTGCGCACAGCTCCCAAGATTTTGGCGATGAACCGCAT 1951
Qy 1933 CATAAAGAACTTCTGACGCTCTGGAAGATCTCATCTGGAATAAGACCTTGAGGCCACT 1992
Db 1952 GATGATCGCCAGCGCAAGTGGCGTTGGATATGGTCTATGATCGCCGACCGAGG---AT 2008
Qy 1993 GAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGAAGAAAGTCAATCAGACT 2052
Db 2009 TACGATCGCTGCGAGGAATTCATGAGCTGTTTGAAGGCGTTTCTGTCGCGATGCCAAG 2068
Qy 2053 GATGAGTGGAAATATGGCCCTGTGGAAGAACGCTTGAATGATGCTTGTGAAGGGCAAT 2112
Db 2069 GATGCTCGCTGAACAGCTGCGCCTATGCTTTGTTTGAAGCTTTGGCACAGCGCATC 2128
Qy 2113 GAAAAACATATTATGAGGATCTAGGAAGCCAGGTTAAACCAAAAAAATATCCCGGA 2172
Db 2129 ATCGACGCGGATAAGAAATGGCCTTTGAGGATGATCTGGAAGCAGGCATGAAGGAGTCT 2188
Qy 2173 CCTCTCAATATAATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTGGTATCTTTT 2232
Db 2189 CCTATTTGGATCATCAAGAGGACCTTCTCAAGCGCATGAAGACCGTGGGTGAGCTGTTT 2248
Qy 2233 GGAGTGGAAAAATATTTTACCTCAGGTTATAAGTCAAGCCGCGGTTTATGAAGAGGCT 2292
Db 2249 GGTTCGGACAGATGACGCTGCCATTCGTGTGCAATCGGCAGAAACCATGAATACTGGG 2308
Qy 2293 GTTGGCCACCTTATCCCTTTTCATGAAATAAGAAAGAAAGAAACACAGAGTCTTTAACGGC 2352
Db 2309 GTGGCCTATTTTGGAACCCGTTCAATGAAGAGGAAGCAGAAAGTACC----- 2353
Qy 2353 ACAGTAGAAGAGGAGGACCTTTACAGGGCACCATCGTCTGGCCACTGTTTAAAGGCGAC 2412
Db 2354 ---GGATCTGGCGAGGAGAGGGCAAGGGCAAAAATCGTCTGGCCACCGTCAAGGGGTGAC 2410
Qy 2413 GTGCACGATAGCAAGAAACATAGTTGGAGTAGTCTTGGCTGCGCAATATTTTCCGAGTT 2472
Db 2411 GTGCACGATATCGGAAGAACTTGTGTGACATCATTTTGTCCAACACGTTACGACGTG 2470
Qy 2473 ATTGATTTAGAGTCACTGACATGATGATTAAGATACCTGAAAGAGTGTCTTTGACACAAA 2532
Db 2471 GTGAACTTTGGGATCATAGCAGCCACTGTGTCGCCCATGTTTGGGAAGCAGCGGAAGAACACAAA 2530
```

```
Qy 2533 GCAGATATAATTGGCTGTTCAGGACT 2558
Db 2531 GCAGCGTCAATCGGCATGTTCGGACT 2556

RESULT 8
US-11-055-822-711
; Sequence 711, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habershauer, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BG1-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 711
; LENGTH: 2578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2578)
; OTHER INFORMATION: PRXA02198
US-11-055-822-711
```

Query Match 2.9%; Score 112.4; DB 12; Length 2578;  
Best Local Similarity 45.5%; Pred. No. 1.5e-20;  
Matches 640; Conservative 0; Mismatches 721; Indels 45; Gaps 5;

```
Qy 1162 CCGTACACCAACTTTGTTTAACATTTGGAGCGCTGTTAATGTTGCAGATCAAGAAAGTTT 1221
Db 1187 CAGAAACCGGCATTTCCATGATCGGTGAGCGCAACCACTCCACGGTTCCAAGGCATTC 1246
Qy 1222 GCTAAACTCATCGGAGGAAACTATGAAGAAGCCTTGTGTGCGAAAGTGCAGGTG 1281
Db 1247 CGTGAGCAATGCTGCTCTGGCGATTGGGAAAGTGTGTGATATTGCAAGCAGCAAAACC 1306
Qy 1282 GAAATGGAGGCCAGGTTGTTGGATGTCAACATGATGATGGCATGTGTAGTGTGTTCCAGT 1341
Db 1307 CGCGATGCTGCACACATGCTGGATCTTTGTGTGATTTACGTGGGACGAGACGGCACCGCC 1366
Qy 1342 GCATGACCAAGATTTTCAACTTAATTTGCTTCCGAGCCAGACATCGCAAGAGTTACCTTTG 1401
Db 1367 GATATGGCGACCTTGGCAGCACTTCTTTGTAC-----CAGCTCCACTTTGCCAATC 1417
Qy 1402 TGCATCGACTCCTCCAATTTTGTGTTGATTGAAGCTGGGTTAAAGTGTCTGCCAAGGGAAG 1461
```

Db	1418	ATGATTGAC	TCCACCGGACGAGAGGTTATTTGCGCAGGCCCTTGAGCACTTGGGTGACGGA	1477
Qy	1462	TGCATTTGTC	CAATAGCATTTAGTCTGAGAGGAAGGAGGACGA-----CTTCTTTGGAG	1512
Db	1478	AGCATCGT	TAACTCCGTCAACTTTGAGACGGCGATGGCCCTGAGTCCCGCTACAGCGC	1537
Qy	1513	AAGCCAGG	AAGATTAATAAGATGAGAGCTGCTATATGTTGTGTCATGCTTTTGTATGAAGAA	1572
Db	1538	ATCATGAA	CTGGTAAACGACGCGTGGCGCGCTGGTTGCGCTGACCAATGATGAGGAA	1597
Qy	1573	GGACAGCA	CAAGAAACAGACACAAAAATCAGAGTGTGCACCGCGGCTTACCATCTGCTT	1632
Db	1598	GGCCAGGC	ACGTATCCCGCTGAGCACAAGGTGGCAATTCCTAAAGCACTGATGTGACGATC	1657
Qy	1633	GTGAAAAA	CTGGGCTTTAATCCAAATGACATTAATTTTGACCCCTAATATCTCTAAACCATT	1692
Db	1658	ACGGCAGC	TACGGCTGGATATCAAAAGACATCGTTGTGACCTGCCTGACCTTCCCGATC	1717
Qy	1693	GGGACTGGA	ATGAGGAAACAACACTTGTATGCCATTAATTTATGCCATGCAACAAAGTC	1752
Db	1718	TCTACTGCC	AGGAAGAAACAGGCGAGATGGCAATCGAAACCATCGAAGCATCCCGCAG	1777
Qy	1753	ATTAAGAA	ACATTTACCTGGAGCGCAGAAATAGTGGAGGTCTTTCCAACTTGTCTCTCTCC	1812
Db	1778	CTGAAGAAG	CTCTACCCAGAAAAATCCACACACCCCTGGGTCTGTCCCAATATTTCTCTCGGC	1837
Qy	1813	TTCCGGAAT	TGGAAGCCATTTGAGAAAGCAATGCATGGGGCTTTTCTTTTACCATGCAATC	1872
Db	1838	CTG-----	AACCTGTGTCAGCCAGGTTCCTAACTCTGTGTCTCTCAATGAGTGCATTT	1891
Qy	1873	AAGTCTGG	CAATGACATGAGAGATAGTGAATGCTGGAAACCTCCCTGTGTATGATGATATC	1932
Db	1892	GAGCTGGT	CTGGACTCTCGGATTTGCGCACAGCTCCAAAGATTTTGGCGATGAACCGCATT	1951
Qy	1933	CATAAGSA	ACTTTCTGACGCTCTGTGAAGATCTCATCTGGAAATTAAGACCCTGAGGCCACT	1992
Db	1952	GATGATCG	CCAGCGAAGTGGCGTTTGGATATGGTCTATGATCGCCGACCCGAGG---AT	2008
Qy	1993	GAGAAGCT	CTTATCGTTATGCCAGACTCAAGGCACAGCGAGGAAGATTCATTCAGACT	2052
Db	2009	TACGATCG	CTGCGAGGAATTCATGACAGCTGTTTGAGGGCGCTTCTGCTGCCGATGCCAAG	2068
Qy	2053	GATGAGTGS	AGAAATGCCCCCTGTGGAAGACGCGCTTGAGTATGCCCTTGGAAGGCGATT	2112
Db	2069	GATGCTCG	CGCTGAACAGCTGGCGCTATGCTCTTTGTTGAGCGTTTGGCACAGCGCATC	2128
Qy	2113	GAAAAACA	TATATTGAGGATCTGAGGAAGCCAGGTTAAACCAAAAAAATATCCCCGA	2172
Db	2129	ATCGACGG	CGGATAGAAATGGCCCTTGAGGATGATCTGGAAGCAGGCAATGAAGGAAGTCT	2188
Qy	2173	CCTCTCAA	TATTAATGAAGACCCCTCATGAAATGGAATGAAATTTGTTGTGTGATCTTTTT	2232
Db	2189	CCTATTGG	ATCATCAACGAGGACCTTCTCAACGCGATGAAGCCGTGGGTGAGCTGTTTT	2248
Qy	2233	GGAGCTGG	AAAAATTTGTTCTACCTCAGGTTATAAAGTCAGCCCGGGTTATGAAGAAGGCT	2292
Db	2249	GTTTCCGG	CACAGATGACGCTGCCATTCGTGCTGCAATCGGCAGAAAAACCATGAAAACTGCG	2308
Qy	2293	GTTGGCCAC	CTTATCCCTTTTCATGGAAAAAGAAAGAGAAAGAACCCAGAGTGCTTTAACGGC	2352
Db	2309	GTGGCCCT	ATTTTGGAACGTTTCATGGAAGAGGAAGCAGAGCTACC-----	2353
Qy	2353	ACAGTAGA	AGAAGAGGACCCCTTACAGGGGCACCATCGTGTGCGCCACTGTTAAAGCGCAC	2412
Db	2354	---GGATCT	CGCGCAGGAGGGCAGGGCAAAATCGTGTGGCACCCGTCAGAGGGTGA	2410
Qy	2413	GTGCAACA	TATAGGCAAGAAACATAGTTGGAGTAGTCTCTGGCTGCAATAATTTCCGAGTT	2472
Db	2411	GTGCACCA	TATCGCAAGAACTTGGTGGACATCATTTTTGTCCAAACACGGTTACGACGTG	2470
Qy	2473	ATTGATTT	AGAGTTCATGCTCCATGTGATAAGATCTGAAAGCTGCTTTTGACCAAAA	2532

[illegible]

```
/ NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 141
; LENGTH: 3477
; TYPE: DNA
; ORGANISM: Thermobifida fusca
US-10-858-730-141

Query Match      2.4%; Score 93.4; DB 8; Length 3477;
Best Local Similarity 43.6%; Pred. No. 4.3e-15;
Matches 628; Conservative 0; Mismatches 766; Indels 45; Gaps 3;

Qy 1145 AGCCCTTCAGGATGGACCGTACCAACTTTTGAACATTTGGAGAGCGCTGTATGTTG 1204
Db 980 AGAGCGTCCCGTTCGCGCAGGACACCGAGTACCTGGCGATCGGGAGCGCAACACGCCA 1039
Qy 1205 CAGGATCAAGGAAGTTTGTCTAACTCATCATGCGAGGAAACTATGAAGAACCTTTGTGTG 1264
Db 1040 ACGGCTCCAGGCGTTCGCGAAGCCATGCTCGCGGAACGCTACGACGACTGTGTGGAGA 1099
Qy 1265 TTGCCAAGTGCAGTGGAAATGGAGCCAGGTTTGGATGTCACATGATGATGGCA 1324
Db 1100 TCGCCCGCCGACAGATCCGCGACGGCGGCACATGCTGACCTGTGGCTGACTATGTGG 1159
Qy 1325 TGCTAGATGTTCCAAAGTGCATGACCAATTTTGCACCTTAATTTGCTTCCGAGCCAGACA 1384
Db 1160 GACGCGAGGGGTGCGGATATCGGGAGCTGGCTTCCGGCTGGCCACCGCTCCACGC 1219
Qy 1385 TCGCAAGGTAACCTTTGTGCAATGACCTCTCCAAATTTTGTGTGATGAAGCTGGGTTAA 1444
Db 1220 TCGCGCTCGTACTGCACTCCACCGAAGTAGCGGTACTGGAAGCTGGACTGGAGATGCTGG 1279
Qy 1445 AGTCTGCTCAAGGGAAGTGCATGTCATATGATGATGTCGTAAGGAAGGAGGACGACT 1504
Db 1280 GCGGCGCGCGTGTCTCAACTCGGTCAACTACGAGGACGGCGACGGCCCTGACTCCCGGT 1339
Qy 1505 TCTTGGAGAGGCCAGGAAGATTAAAGATGATGAGCTGCTATGTTGTGTCATGCTTTTG 1564
Db 1340 TCGCAAGTTCGCGCGTGGCGGTGGAGCAGCGGGCGGCCCTCATGGCGTGAACCATCG 1399
Qy 1565 ATGAAGAAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGACCCGGCGCTTACC 1624
Db 1400 ACAGACAGGGGACGGCGGCGGACCGCGAAAGTGGAGTGGCGAGGCTGCTCATCC 1459
Qy 1625 ATCTGCTTGTGAAAAAATCGGCTTTAATCCAAATGACATTTATTTGACCCCTAATACC 1684
Db 1460 GGCACCTCACCACCGAGTACGGCATCCGACAGCAGCATCATCTGTGAGCTGCGCTGACCT 1519
Qy 1685 TAACCATTTGGACTGGAAATGGAGGAACACAACTTGTATGCCATTAATTTATCCATGCAA 1744
Db 1520 TCACGATCGCAACCGACAGAGGAGTTCGGCGGCGAGCTCTGGAACACCATCGAGGGA 1579
Qy 1745 CAAAGTCATTAAAGAAACAATTACTGGAGCCAGAAATAGTGGAGTCTTTCCAACTTGT 1804
Db 1580 TCCGTGAATGAAGCGGCGCCACCGGACGTGACAGACCGCTGGCGCTGTGCTCCAA----- 1634
Qy 1805 CTTCTCTCTTCGAGGAATGGAAGCATTCGGAAGCAATGATCGATGGGTTTTCCTTTACC 1864
Db 1635 -CGTCTCTCTTCGAGTCAACCGGCTGCCCGCATTTGTCTCAACTCGGTGTTCCTCCACG 1693
Qy 1865 ATGCAATCAAGTCTGGCATGACATGAGATGAAATGCTGGAAACCTCCCTGTGTATG 1924
Db 1694 AGTGGTTCAGCGCGCTTGGACTCCGCGATCGTGACGCTCCAGATCTCGCGCATCA 1753
Qy 1925 ATGATATCAATAGGAACTTCTGCACTCTGTGAAGATCTCATCTGGAATTAAGACCCCTG 1984
Db 1754 ACCGCATCCCGAGGAGCAGCGGAGGTGGGCTTGGACATGATCTACCGCGCGCACCG 1813
Qy 1985 AGGCACCTGAGAGCTCTTATGTTATGCCAGACTCAAGGCACAGGAGGAAGATCA 2044
Db 1814 ATGACTACGCCGCTGCAAGCGTTCTCGACGCTTTTCAAGGAGTGGACGCGCAGGGA 1873
Qy 2045 TTCAGACTGATGAGTGGAGAAATGCGCCTGTGGAAGAACGCGCTTGTGATGCGCCTTTGTA 2104

1874 T---GCGCGCTCGCGCGAGGAAGAGCTGGCGCGCTGCGCTGTGTGGAGCGCTGGAGC 1930
2105 AGGCATTGAAACACATATTATTGAGGATACTGAGGAAGCCAGGTTAAACCAAAAAAAT 2164
1931 GCCGTATCGTCGACGGGGAAGCCGCCGCGATGGAAGCGGACCTGGAGCAAGCGCTCACC 1990
2165 ATCCCCGACCTCTCAATATATTAATTGAAGGACCCCTGATGAATGGAATGAAATTTTGGTG 2224
1991 AGCGTCCGCGCTGGACATCATCAACACACGCTGCTGGCGGGATGAAGACCGCTCGCG 2050
2225 ATCTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTATAAGTTCAGCCGGGTTATGA 2284
2051 ACCTGTTGGCTCGCGGACAGTCCGCTCCGCTGCTGCTGAAGTTCGCGGAGTGATGA 2110
2285 AGAAGCTGTGGCCACTTATCCCTTTTCATGNAAGAAAGAGAGAAACACAGAGTGC 2344
2111 AGCGCGCGTGGCTTATCTGGAGCCGACATGGAAGAGGTG----- 2151
2345 TTAACGCGCACAGTAGAAGAAAGAGACCCCTTACAGGGGACACATCGTCTGCGCCACTGTA 2404
2152 -----GACGGCGACCTCGGCAAGGGCGGATCGTGTGCGCACCGGTCA 2194
2405 AAGCGCAGCTGCACGACATAGGCAAGACATAGTTGGAGTAGTCTTGGCTGCAATAATT 2464
2195 AGGCGCAGCTCCACGACATCGCAAGAACCTTGTGGACATCATCTGTCCAACACCGCT 2254
2465 TCCGAGTTATGTTTAGGAGTCATGACTCCATGATAGATAGTACTGAAAGCTGCTCTTG 2524
2255 ACGAGGTCATCAACCTGGGATCAAGCAGGCCCATCTCCGCGATCTCGAGGCGGCCGAGC 2314
2525 ACCACAAAGCAGATATTAATTGGCTGTGAGACTCATCACTCTCTTCCCTGGATGAAATG 2583
2315 GGCACCGCGCGAGCTGATCGGCATGTCGCGCTGCTGCTGAGTCCACGCTGCTGATG 2373

RESULT 11
US-10-858-730-142
; Sequence 142, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142
; LENGTH: 3513
; TYPE: DNA
; ORGANISM: Streptomyces coelicolor
US-10-858-730-142

Query Match      2.3%; Score 91.4; DB 8; Length 3513;
Best Local Similarity 44.3%; Pred. No. 1.6e-14;
Matches 642; Conservative 0; Mismatches 741; Indels 66; Gaps 4;

Qy 1145 AGCCCTTCAGGATGGACCGTACCAACTTTTGAACATTTGGAGAGCGCTGTATGTTG 1204
```

Db 1019 AGACCGTGCCCTTCGCGCAGGACACCTCCTACCTGGCCATCGGCGAGCGCACCAACGCCA 1078  
Qy 1205 CAGGATCAAGGAAGTTTGTAACTCATCATGCGCAGGAACATATGAAGAACCTTGTGTG 1264  
Db 1079 ACGGGTCCAAGAGTTTCGCGAGGCCATGCTGGACGGCGCTGGGACGACTGGGTGAGA 1138  
Qy 1265 TTGCCAAGTGCAGGTGAATGGAGGCCAGCTGTTGGATGTCAACATGATGATGGCA 1324  
Db 1139 TGGCCCGCGACCAAGTCCGGAAGGCGCGCATGCTCGACCTTGGGTGACTACGTCG 1198  
Qy 1325 TGCTAGATGGTCCAAGTGCATGACAGATTTTCCAACTTAATTTGCTTCCGAGCCAGACA 1384  
Db 1199 GCCGGGAGCGGCTCGCGACATGAGGAACTGGCCCGCGCGTTCGCCACGCCCTCCAGC 1258  
Qy 1385 TCGCAAGAGTACCTTTGTGATCGACTCCTCCAAATTTTGTGCTGATTAAGAGCTGGGTAA 1444  
Db 1259 TGCCGATCGTCTCGACTCCACCGAGTTCGAGTTCATCCGCGCGCGCTGGAGAGCTCG 1318  
Qy 1445 AGTCTGCGCAAGGAAGTGTCAATGCTCAATAGCATTTAGTCTGAAGAGAGGAGAGCACT 1504  
Db 1319 GCGGCCGCGGGTGTCAACTCGGTCAACTACGAGGACGGCGCGCGCGGAGTCCCGGT 1378  
Qy 1505 TCTTGGAGAGGCGCAGAGAGATTAAAGATAGAGTGTGCTGTATGCTGTATGCTTTTG 1564  
Db 1379 TCGCCCGGCTCAGAAAGTCCCGCGGAGCAGCGCGCGCGCTGATCGCGTGAACATCG 1438  
Qy 1565 ATGAAGAGGACAGGCAACAGAAACAGACACAGAAATCAGAGTGTGCACCGCGGCTTACC 1624  
Db 1439 ACGAGTGGGACAGCGCGCACCGCGAGAGAGTTCGAGATCGCGACGCGTCAATCG 1498  
Qy 1625 ATCTGCTTGTGAAAAAATCGGCTTTAATCAAAATGACATTAATTTTGGACCTTAATTC 1684  
Db 1499 ACGACCTCACCGCAACTGGGGCATCCACGAGTCCGACATCTCGTGCAGTGCCTGACCT 1558  
Qy 1685 TAACATGGGACTGGATGAGGAGAACACACTTTGATGCAATTAATTTTATCATCAACAA 1744  
Db 1559 TCACCATCTGACCGCGCAGGAGGATCCCGCAAGGACGGCTGGCCACCACCATCGAGGCA 1618  
Qy 1745 CAAAAGTCATTAAAGAAACATTACTCGAGGACCAATAAGTGGAGGTCTTTTCCAACATGT 1804  
Db 1619 TCGGGAACTCAAGCGGGCCACCGAGCGTGACACCGCTCGGCTGTGCAATCT 1678  
Qy 1805 CTTTCTCTTCGAGGAATGGAAGCCATTTCGAAAGCAATGCAATGCGGCTTTCTTTTACC 1864  
Db 1679 CTTTGGGCTC-----AACC CGCGCGCGCATCTCTGCTCAACTCGCTTCTCTCGACG 1732  
Qy 1865 ATGCAATCAAGTCTGGCATGACATGAGATAGTGAATGCTGGAACCTCCTCTGTATG 1924  
Db 1733 AATCGCTCAAGCGCGGCTGACTCGGCCATCGTGCAGCGAGCAAGATCTGCGCGATCG 1792  
Qy 1925 ATGATATCCATAAGGAACCTTCGACGCTCTGTGAAGATCTCATCTGGAATAAAGACCTG 1984  
Db 1793 CCGCTTCGACGAGGAGAGTCAACCGCGCTCGACTTGATCTAGACCGCGCGCGCG 1852  
Qy 1985 AGG-----CCACTGAGAAGCTCTTACGTTATGCCCGAGACTCAAGGACAGGAGGGA 2035  
Db 1853 AGGCTACGACCCCTCGAAAAGCTCATGACGCTCTTCGAGGGCGCCACCGCCAAAGTCG 1912  
Qy 2036 AGAAGTCATTACAGATGATGAGTGGAGAAATGCGCCCTGTGGAAGAGCGCTTGAATG 2095  
Db 1913 TGAAGGCTCCAAAGCGCGGAGAACTGGCGCGCTTCCGCTGGAGGAGCGCTCAAGCGCC 1972  
Qy 2096 CCCTTGTGAAGGGCATTTGAAAAACATATTATTGAGGATACCTGAGGAAGCGAGTTAAACC 2155  
Db 1973 GCATCATGACGCGGAGAGAAACCGGCTTCGAAACAGGACCTCGACGAGGCC----- 2022  
Qy 2156 AAAAAAATATCCCGACCTCTCAATATAATTGAAGGACCCCTGATGAATGAATGAATA 2215  
Db 2023 --CTCCGGGAGCGCGCGCTCTGAGATCGTCAACGACACCTCTCTCGACGGTATGAGG 2080  
Qy 2216 TTGTTGTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATTAAGTCAAGCC 2275

Db 2081 TCCTCGCGAGCTTTCGGCTCGGCGCAGATGAGCTGCGTTCTGCTCAGTCCGCG 2140  
Qy 2276 GGTTTATGAAGAGGTGTTGGCCACTTATCCCTTTTCGAAAAAGAAAGAGAA 2335  
Db 2141 AGGTCATGAAGACCGCGTGGCCCACTTGGAGCGCAGCATGAGAGAGACCGACG 2200  
Qy 2336 CAGAGTGTCTTAACGGCACAGTAGAAGAGAGGACCTTACCAAGGACCATCTGCTGG 2395  
Db 2201 GCA-----AGGGCAGCATCTGCTGG 2221  
Qy 2396 CCACTGTAAAGCGAGCTGCAGCAGATAGGCAAGAACATAGTTGGAGTAGTCTTGGCT 2455  
Db 2222 CCACGCTCGCGGAGCTCCACGACATCGCAAGAACCTCTCGTGCATCATCTGTCCA 2281  
Qy 2456 GCAATATTTCCGAGTTATTGATTTAGGAGTCAATGACTCATGTGATAAGATCTGAAG 2515  
Db 2282 ACAACGCTACAAAGCTGCTCAACCTCGGCATCAAGCAGCGCTCTCGCGATCTCGAAG 2341  
Qy 2516 CTGCTCTTGACCAAAAGAGATATAATTTGGCTGTGAGGACTCATCACTCTTCCCTGG 2575  
Db 2342 CGCCGACAGGACCGGCGCGCTCATCGGCATGTCGGCTCTCTCGTCAAGTCCACGG 2401  
Qy 2576 ATGAAATGA 2584  
Db 2402 TGATCATGA 2410

## RESULT 12

US-10-858-730-144  
; Sequence 144, Application US/10858730  
; Publication No. US2005025588A1  
; GENERAL INFORMATION:  
; APPLICANT: Bailey, Richard B.  
; APPLICANT: Blomquist, Paul  
; APPLICANT: Doten, Reed  
; APPLICANT: Driggers, Edward M.  
; APPLICANT: Madden, Kevin T.  
; APPLICANT: O'Leary, Jessica  
; APPLICANT: O'Toole, George  
; APPLICANT: Trueheart, Joshua  
; APPLICANT: Walbridge, Michael J.  
; APPLICANT: Yorgey, Peter S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID  
; FILE OF INVENTION: PRODUCTION  
; FILE REFERENCE: 14184-030001  
; CURRENT APPLICATION NUMBER: US/10/858,730  
; CURRENT FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/475,000  
; PRIOR FILING DATE: 2003-05-30  
; PRIOR APPLICATION NUMBER: US 60/551,860  
; PRIOR FILING DATE: 2004-03-10  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 144  
; LENGTH: 3621  
; TYPE: DNA  
; ORGANISM: Mycobacterium leprae  
US-10-858-730-144

Query Match 2.3%; Score 89.6; DB 8; Length 3621;  
Best Local Similarity 44.1%; Pred. No. 5.2e-14;  
Matches 1039; Conservative 0; Mismatches 1199; Indels 120; Gaps 11;

Qy 234 AGATCATGCCAGCGCTGAAAGGCAACATGATTTTAAGTATTAATCAGCTGATGT 293  
Db 123 AGATGATTTCCGGGCTGGAGGCTGCAACAGAGATTTCAAGAAACGGCTCTGACGT 182  
Qy 294 CATTTACCAATCCATAAGGAATACTTGTGCTGGGGCAGATATCATTTGAAACAATAC 353  
Db 183 GCTGGAACCATCCACCGACGCTACTTCGAGGAGGTGCGGACCTCTGCGAGACCAAC 242  
Qy 354 TTTAGCAGCACTAGTATTATCCCAAGCTGACTTGGCTTGAACACTTGGCCTTACCGAT 413

Db 243 TTTCGGCTGCAACCTGTCCAACTTTGGTGACTACGACATCGCCGACAAGATCAGGACTT 302  
Qy 414 GAAATGTGCTCTGCGAGGAGTGGCCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGG 473  
Db 303 GTCGACGCGGGCACCGGTGATTTGGCGACGGGTGCGCGACGAGCTGACCAACC---CCGA 359  
Qy 474 AATTAAGAGGTTTGTGGCAGGGGCTCTGGGTCCGACTTAATAAGACACTCTCTGTGTGCC 533  
Db 360 CCACAGCGATACGTCTGGGGTGTGATGGGACAGGACCAAGTTGGCCACCTTGGGCA 419  
Qy 534 ATCTGTGGAAGGCGCGATTAAGAAATACACATTCATTTGATGAGCTTGTGTAAGCATACCA 593  
Db 420 CACCGAGTACCGG-----GTGCTTCGAGACGCGCTACAC 452  
Qy 594 AGACGAGCCAAAGGACTTCTGGATGGCGGGTGTATCTTACTATGAACTATTTT 653  
Db 453 CGATCGCGCTTGAAGCATGCTGGACGCGTGGCTGACGCCGTACTGGTTGAACCTGTCA 512  
Qy 654 TGATACTGCCAATGCAAGGAGCGCTTGTTTGCACTCCAAAATCTTTTGAAGGAGAATA 713  
Db 513 GGACTTGTCTGAGCTCAAGGCTCGGTGCTGGGCTCGCGCGCGATGACACAGGCGG 572  
Qy 714 TGCTCCCGGCTATCTTTATTTTCAAGGACGATCGTTGATAAAGTGGGCGGACTCTTTTC 773  
Db 573 TCGGCACATTCGGGTCTTCGTCCAGTGACTGTGAGACGACCGGAAACGATGCTGCTGG 632  
Qy 774 CGGACAGACGAGGAGGATTTGTCACTCAGCGTGTCTCATGAGAAACCACTCTGCATTGG 833  
Db 633 AAGTGAGATCGCGCTGCACTGGCTGCGCTCGAGCGCTCGGTGTC---GACATGATCGG 689  
Qy 834 ATTAATTTGCTTTGGGTGAGCTGAGATGAGACCTTTTATGAAATAATTTGAAATAAG 893  
Db 690 TTTGAACTGCGAACGGGCGCTGAGATGAGTGAAGCATCTGGCGCACTTGTCCAAGCA 749  
Qy 894 TACAACAGCCTATGTCTCTGTATTCCAATGCAAGTCTTCCCAACACCTTTTGTGACTA 953  
Db 750 TGCCCGCATCCGGTGTCCGTGATGCCCAACGCGGGCTGCGGTGCTGGGTGCCAAGG 809  
Qy 954 TGATGAACGCTTCTATGATGGCAAGCACTTAAGGATTTTGTATGATGCTTGGT 1013  
Db 810 AGCTGAATACCGCTGCAAGCCGCAAGATTTGGCGGAAGCTTTGGCTGGTTTCATCGCTGA 869  
Qy 1014 CAAT-----ATAGTTGAGGATGCTGTGGGTCAACACAGATCATATCAGGAAAT 1064  
Db 870 ATTTGGTCTTTCGTTGGTGTGCTGTGTGATGACACCCCGACCACTCCGGAAGT 929  
Qy 1065 TGCTGAAGCTGTGAATAATTTAAGCCTAGAGTTTCCACTCGCACTGCTTTTGAAGGACA 1124  
Db 930 GCGCGAGCGGTAGCCAGATGCAACGAGCGGACAGTGCCACGCGGTGAGCGTCAATGTGAC 989  
Qy 1125 TATGTTACTGTGTGTAGAGCCCTTCAGATTTGGACCGTACACCACT-----T 1175  
Db 990 CTATGAGCGCTCGGTATCGTCTGTATACAGCCATTCATTTGCGCCCAAAACCCCTCGGT 1049  
Qy 1176 TGTTAACTATGAGAGCGCTGTAAATTTGCGAGGATCAAGGAAGTTTGTAAACTCATCAT 1235  
Db 1050 TCTGATGATCGGTGAGCGTGAAGATGCCAACGCGTCCAAAGTTTTCGTGAGGCAATGAT 1109  
Qy 1236 GCGAGGAACATTAAGAAAGCCTTGTGTGTGCCAAAGTGCAGGTGGAATTTGGAGGCCA 1295  
Db 1110 GCGCGAGGACTATCAAAAGTGTAGATATGCGCAAGGACCAAAACCCGTGGCGGCGACA 1169  
Qy 1296 GGTGTTGATGTCAACATGATGATGGCATGTGTAGATGCTCAAGTGCATGACACAGATT 1355  
Db 1170 CCGTGTGATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229  
Qy 1356 TTGCAACTTAATTTGCTCCGAGCGACATCGCAAGGTTACCTTTGTGATCGACTCTCTC 1415  
Db 1230 GCGCGGTGCTTGCACCGTGTGACATTTGCCGATCATGC-----TGACTCTAC 1280  
Qy 1416 CAATTTTGTGTGATTTGAAGCTGGGTAAAGTGTGCCAAGGAGTGTGATTTGTCAATAG 1475  
Db 1281 CGAATACCGGTGCTGCGAGGAGGTTTGGAGCACCTCGGCGGCGCTCGGTGATCAAATTC 1340

Qy 1476 CATTGTCTGAAGGAGGAGGACGACTTCTTTGGAGAGCGCCAGGAGATTAA----- 1529  
Db 1341 CGTCAACTACGAGACGCTGACCGTCCGAGTCAAGTTCGAGTCAAGACCACTGGAGCTGGT 1400  
Qy 1530 ---AAAGTATGGAGCTGCTATGGTGGTTCATGGCTTTTGTATGAAGAAAGGACAGCAACAGA 1586  
Db 1401 GCGCGAGCACGAGCGCGGTGGTGGCTGACCATCGACGAAACAGGCTCAGGCCCGCAC 1460  
Qy 1587 AACAGACAAAAATCAGAGTGTGCACCGGGCTTACCATCTGCTTGTGAAAAAATCTGGG 1646  
Db 1461 CGTTGAAGAAGGTGGAAGTCCGAGCGGCTTATCAATGACATTAACGATTAACCTGGG 1520  
Qy 1647 CTTTAATCCAAATGACATTAATTTTGAACCCCTAATATCTTAACCATTTGGGACTGGAATGGA 1706  
Db 1521 CGTTGATAAATCGCGGATTTCTCATGATGCTTGTGACTTTTACTATTGGCACTGGCCAGGA 1580  
Qy 1707 GGAACACAACTTGTATGCCATTAATTTTATCCATGCAACAAAAAGTCAATTAAGAAAAAT 1766  
Db 1581 GGAGTCACGCAAAAGACGGCATTCGAGACCATCGACGCGATTCGTGAGCTGAAGAAGCGCA 1640  
Qy 1767 ACCTGGAGCCAGATAAGTGGAGGCTTTTCCAACTTGTCTTCTCTTCCGAGGAATGGA 1826  
Db 1641 CCAGCGGTGACACTACGCTGGGTTGTCCAA-----CATCTCTTCTCGGTCTCAATCC 1694  
Qy 1827 AGCCATTTGAGAGAACATGCAATGGGGTTTCTTTCATGCAATCAAGTCTGGCATGGA 1886  
Db 1695 TTCTGACGCCAAGTCTTAACTCTGTGTTTTCTACATGAATGTCCGAGACGAGCATGGA 1754  
Qy 1887 CATGGAGATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATCCATAGGAACCTTCT 1946  
Db 1755 TTCCGCGATTTGACAGCTTCAAGATATTGCCCATCAACCGGATACCCGAGAACAGCG 1814  
Qy 1947 GCACTCTGTGAAGATCTCATCTGGAAT-----AAAGACCTGAGGCCACTGAGAA 1997  
Db 1815 CCAGGCTGCGCTGATCTAGTGTATGACCGCGCTCGCGAAGGTTACGACCATTTGCGAGAA 1874  
Qy 1998 GCTCTTACGTTATGCCAGACTCAAGGACAGAGGAGGAGAAAGTCAATTCAGACTGATGA 2057  
Db 1875 GCTGATGTGTTATTAAGAGTGTGCTGCCATCTGTCGAGGAACACACGAGGAGGAGA 1934  
Qy 2058 GTGAGAAATTTGGCCCTGTGCAAGAACGCTTTGAGTATGCCCTTTGTGAAGGCGCATGAAA 2117  
Db 1935 ACTCGCTAAGCTCGGTGTTGTCACCGTTAGCACAGCGGATCGTCGACGCGAGCGCAA 1994  
Qy 2118 ACATATTAATTTGAGGATACTCAGGAAGCAGGTTTAAACCAAAAAAATAATCCCGACCTCT 2177  
Db 1995 CGGTTAGATGTTGATCTCGACGAGGCAATGACCCAGAAA-----CGCGCGTT 2042  
Qy 2178 CAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTTGGTGTATCTTTTGGAGC 2237  
Db 2043 GCGCATCATCAACGAGAACCTTGTGACGCGCATGAGACAGTCTGCTGAATTTGTTGCGCTC 2102  
Qy 2238 TGGAATAATTTTCTACTCAGGTTTAAAGTTCAGCCCGGTTATGAAGAGGCTGTGG 2297  
Db 2103 TGCGCAGATGCGCTGCTTTCGTTGTCAGTCCGCGAGGTTATGAAGACGAGCGGTGGC 2162  
Qy 2298 CCACCTTATCCCTTTTCAATGGAATAAGAAAGAAAGAACCAAGTGTCTTAAGGCAAGT 2357  
Db 2163 TTATCTAGAACCGCACATGGAGAAA-----TCCGACTGTGA 2198  
Qy 2358 AGAAGAAGAGGACCCCTTACGAGGCGCACCATCTGCTGCGCACTGTTAAAGCGGACGTGCA 2417  
Db 2199 CTTCCGTAAGGGGTTAGCAAGAGCGGATTTGCTGCTACCGTCAAGAGAGATGTGCA 2258  
Qy 2418 CGACATAGGCAAGAAACATAGTTGAGTAGTCTTGGCTGCAATTAATTTCCGAGTTATTGA 2477  
Db 2259 CGATATTTGGCAAAACCTCTGTCGATATCATTTCTGAGCAACAAACGCTACGAGTGTGAAA 2318  
Qy 2478 TTTAGGAGTCAATGCTCAATGTGTATGAAGTCTGTGAAAGCTGCTCTTGAACCAAGAGAGA 2537  
Db 2319 CCTCGCATCAAGCAGCGGATTTACCAACATTTCTGAGGTGGCCGAGGACAAAAGCGCGA 2378





Db 2342 CCGACGTGGTGGCATGTGCGGCGCTGCTGGTGAAGTCGACCGTGGTGATGAAGGAAACC 2401  
QY 2594 CCAAGGAATGGAGA 2608  
Db 2402 TCGAGGAGATGAACA 2416

## RESULT 14

US-11-175-859-4373  
; Sequence 4373, Application US/11175859  
; Publication No. US20060024715A1  
; GENERAL INFORMATION:  
; APPLICANT: Affymetrix, Inc.  
; TITLE OF INVENTION: Method of Analysis of Human Polymorphism  
; FILE REFERENCE: 3690.1  
; CURRENT APPLICATION NUMBER: US/11/175,859  
; CURRENT FILING DATE: 2005-07-05  
; PRIOR APPLICATION NUMBER: US 60/585,352  
; PRIOR FILING DATE: 2004-07-02  
; NUMBER OF SEQ ID NOS: 116251  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4373  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: homo sapien  
US-11-175-859-4373

Query Match 1.2%; Score 46.4; DB 12; Length 50;  
Best Local Similarity 94.0%; Pred. No. 0.0052;  
Matches 47; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2794 GAAGAATATGAAGATATTAGACAGACCATTATGAGTCTCTCAAGGAGAG 2843  
Db 1 GAAGAATATGAAGATATTAGACAGGCCATTATGAGTCTCTCAAGGTAAG 50

## RESULT 15

US-09-925-065A-483679/c  
; Sequence 483679, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 483679  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-483679

Query Match 1.1%; Score 43.2; DB 6; Length 616;  
Best Local Similarity 54.4%; Pred. No. 0.21;  
Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 3721 TTCTCAATTTCAAGTCCTCAAAATATTTTGTGTGGGGAAGATTTCCAAGGATCAGGTTGAG 3780

Db 302 TTCACACATTTTTTTTCCAAAGTAATAATTTGGAATACTTTTTATCTCAACACCAAGTATTAG 243  
QY 3781 GATTATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCAATT 3840  
Db 242 GATTTTGTGTTGGAAATTAATATATCTATAGATAAAATTTAGGGGTCTCTTGGTAACACT 183  
QY 3841 TTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGC 3880  
Db 182 ATTGATTCTCATACAGACTATGCTAGTTCTATGTTCTTAC 143

Search completed: March 6, 2006, 19:47:06  
Job time : 713 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 01:52:43 ; Search time 13579 Seconds  
(without alignments)

13503.099 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggcacactgtggagagcagc.....ctcaaggaataacaactag 3919

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 18

Total number of hits satisfying chosen parameters: 312259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_est8.\*

9: gb\_est9.\*

10: gb\_gss2.\*

11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2460	62.8	3711	11	DQ037675 Homo sapi
2	839	21.4	4888	4	CR859882 Pongo pyg
3	681	17.4	863	3	EM450887 AGENCOURT
4	660	16.8	721	1	AL598311 DKEZp313F
5	651	16.6	871	5	EX390174 EX390174
6	637	16.3	943	5	BU189680 AGENCOURT
7	623	15.9	698	7	CU414972 170004180
8	545	13.9	583	3	BF313585 BP313585
9	529	13.5	582	3	BP336553 BP336553
10	501	12.8	583	3	BP337383 BP337383
11	497	12.7	560	3	BP220649 BP220649
12	490	12.5	490	1	AL041796 DKEZp434B
13	490	12.5	3677	11	DQ037676 Pan trogl
14	480	12.2	582	3	BP336015 BP336015
15	448	11.4	1043	3	BQ062840 AGENCOURT
16	440	11.2	626	5	EX485695 DKEZp686C
17	404	10.3	404	7	CV334327 IL3-NT028
18	363	9.3	937	5	BQ960377 AGENCOURT
19	350	8.9	874	2	BE879762 601491666
20	332	8.5	579	3	BP268350 BP268350
21	323	8.2	527	5	EX470290 DKEZp779J
22	321	8.2	523	1	AW277117 xp63a10.x

23	315	8.0	506	3	BQ364647
24	315	8.0	937	2	BG258321
25	314	8.0	961	5	BX332266
26	298	7.6	574	7	CV344680
27	287	7.3	581	3	BP336596
28	267	6.8	587	5	BX498595
29	267	6.8	587	5	BX498595
30	251	6.4	583	3	BP332632
31	236	6.0	466	5	BX484356
32	232	5.9	581	3	BP348983
33	231	5.9	592	3	BP249465
34	216	5.5	496	5	BU429497
35	212	5.4	618	7	CV356711
36	205	5.2	232	1	AA356439
37	201	5.1	356	5	BX955050
38	180	4.6	581	3	BP248483
39	158	4.0	996	5	BU184298
40	154	3.9	668	1	AI872333
41	142	3.6	477	7	CV363529
42	136	3.5	552	7	CR767594
43	134	3.4	226	2	BF749082
44	129	3.3	486	7	CR773515
45	128	3.3	444	7	CR751928
46	127	3.2	1168	5	BQ960236
47	124	3.2	289	7	CR853354
48	117	3.0	857	9	AQ744889
49	114	2.9	735	10	AG092782
50	98	2.5	153	6	CF135946
51	98	2.5	436	1	AU077264
52	98	2.5	517	3	BQ311389
53	98	2.5	554	3	BP380788
54	96	2.4	599	6	CD675851
55	96	2.4	623	8	DN996221
56	96	2.4	639	7	CN414973
57	88	2.2	550	3	BP380920
58	86	2.2	252	2	BI027486
59	81	2.1	607	7	CR770704
60	68	1.7	581	8	DN279755
61	68	1.7	715	8	DN280822
62	60	1.5	551	9	AQ411010
63	57	1.5	248	2	BE929041
64	54	1.4	217	10	CL211992
65	54	1.4	458	7	CK389585
66	54	1.4	539	6	CF165176
67	54	1.4	658	6	CB455191
68	54	1.4	727	6	CB600441
69	54	1.4	1326	10	AG448662
70	52	1.3	477	7	CV363529
71	51	1.3	2150	4	AK037599
72	50	1.3	162	2	BF852392
73	50	1.3	538	6	CB454749
74	47	1.2	365	7	CK465144
75	47	1.2	365	7	CK467246
76	46	1.2	396	2	BF515662
77	43	1.1	445	7	CR751294
78	41	1.0	544	1	AJ448453
79	41	1.0	648	1	AJ724456
80	41	1.0	717	1	AJ724468
81	40	1.0	452	7	CN690949
82	39	1.0	520	3	BI345164
83	39	1.0	571	3	BI345087
84	38	1.0	548	2	BE684351
85	38	1.0	617	7	CK620511
86	38	1.0	633	2	BB662292
87	38	1.0	678	1	BB039567
88	38	1.0	758	8	CK228713
89	38	1.0	909	2	BG918727
90	38	1.0	915	5	BQ954233
91	38	1.0	954	6	CA786770
92	38	1.0	1788	4	AK085987
93	37	0.9	156	1	AA285223
94	37	0.9	501	3	BM088286
95	37	0.9	809	9	BZ873476

96	35	0.9	397	2	BB813231	BB813231	169	26	0.7	656	6	CB528401	CB528401	UI-H-FT2-
c 97	35	0.9	822	7	CK846954	969526 MA	170	26	0.7	658	2	BB657977	BB657977	BB657977
c 98	34	0.9	292	2	BE241066	BE241066	171	26	0.7	708	6	CA763245	CA763245	CA763245
c 99	34	0.9	389	9	AZ485259	IM0312N03	172	26	0.7	725	5	BQ780303	UI-R-RFO-	BQ780303
c 100	32	0.8	688	10	CE401545	tigr-g88-	173	26	0.7	727	3	CR433139	CR433139	CR433139
c 101	32	0.8	798	10	CG843722	Yhhw4336	174	26	0.7	746	3	BQ003913	UI-H-E11-	BQ003913
c 102	30	0.8	591	5	BX510153	DKF2p686H	175	26	0.7	752	7	CV598811	L AH-aaa3	CV598811
c 103	31	0.8	402	3	BQ365637	MR2-GN003	c 176	26	0.7	757	1	AJ724469	AJ724469	AJ724469
c 104	30	0.8	566	10	CE321611	tigr-g88-	177	26	0.7	778	6	CA307248	UI-H-FT1-	CA307248
c 105	30	0.8	698	5	BW782486	BW782486	178	26	0.7	806	1	AJ724455	AJ724455	AJ724455
c 106	29	0.7	363	1	AI604746	VC20d02.Y	179	26	0.7	812	5	BX752969	BX752969	BX752969
c 107	29	0.7	398	6	CA531346	CS0318D05-	180	26	0.7	826	7	CK197521	FGAS00599	CK197521
c 108	29	0.7	436	1	AA387056	VC20d02.r	181	26	0.7	1094	11	AL322116	Tetrarodon	AL322116
c 109	29	0.7	486	9	AZ765154	IM0561M20	c 182	26	0.7	1105	9	CC184546	CH261-52N	CC184546
c 110	29	0.7	551	10	CL352328	RPC144.40	c 183	25	0.6	78	7	CN545963	EST 17911	CN545963
c 111	29	0.7	589	2	B1065590	pgfin.pk0	c 184	25	0.6	95	10	BX120837	Danio rer	BX120837
c 112	29	0.7	613	1	AL872457	AL872457	c 185	25	0.6	102	6	CB376734	HD06D03.L	CB376734
c 113	29	0.7	627	1	AL875075	AL875075	c 186	25	0.6	103	6	CF510960	CABud0001	CF510960
c 114	29	0.7	657	9	CC249505	RRF186.Ba	c 187	25	0.6	105	6	CF215276	CAS70001	CF215276
c 115	29	0.7	665	1	B1156044	B1156044	c 188	25	0.6	109	1	AI340607	tb31h02.x	AI340607
c 116	29	0.7	812	7	CF575710	CF575710	c 189	25	0.6	110	6	CF269835	Fcylcolid1	CF269835
c 117	29	0.7	885	6	CF548421	AGENCOURT	c 190	25	0.6	110	7	CN385319	LE2TR03E1	CN385319
c 118	28	0.7	161	3	B1376557	BFLG3.000	c 191	25	0.6	115	7	CN545789	EST 17733	CN545789
c 119	28	0.7	225	2	BE005215	CM1-BN011	c 192	25	0.6	117	7	CN545605	EST 17549	CN545605
c 120	28	0.7	230	9	AZ844485	2M0143P24	c 193	25	0.6	117	7	CN546348	EST 18300	CN546348
c 121	28	0.7	518	7	CO596140	DG8-121j5	c 194	25	0.6	118	6	CF518976	CAP0006.I	CF518976
c 122	28	0.7	622	7	CN537801	UI-M-HS0-	c 195	25	0.6	124	6	CF512448	CABud0003	CF512448
c 123	28	0.7	669	1	AJ724454	AJ724454	c 196	25	0.6	127	7	CN546260	EST 18212	CN546260
c 124	28	0.7	693	5	BU290174	604161924	c 197	25	0.6	129	6	CF212016	CGF100066	CF212016
c 125	28	0.7	717	6	CF739229	UI-M-HD0-	c 198	25	0.6	129	6	CF212499	CGF100066	CF212499
c 126	28	0.7	728	5	BV372990	603811666	c 199	25	0.6	132	6	CB978871	CAB40006	CB978871
c 127	28	0.7	738	1	AJ724471	AJ724471	c 200	25	0.6	132	8	DN837912	SmoC-1.01	DN837912
c 128	28	0.7	758	10	AG546553	MuB.M8CU	c 201	25	0.6	136	1	AI275354	qw66e02.x	AI275354
c 129	28	0.7	914	5	B0343179	603524953	c 202	25	0.6	136	6	CF518719	CAR0007.I	CF518719
c 130	27	0.7	101	8	CV994689	IPCGR3.8	c 203	25	0.6	143	5	B0531347	AGENCOURT	B0531347
c 131	27	0.7	263	2	B7333369	MR1-AN003	c 204	25	0.6	143	5	B0531590	ALBEDO000	B0531590
c 132	27	0.7	421	1	AJ825014	AJ825014	c 205	25	0.6	143	6	CB610600	AGENCOURT	CB610600
c 133	27	0.7	453	6	CB077007	h147e10.g	c 206	25	0.6	143	6	CF205875	RR890915I	CF205875
c 134	27	0.7	510	10	CL336386	RPC144.26	c 207	25	0.6	144	5	B0563912	AGENCOURT	B0563912
c 135	27	0.7	579	10	CL321842	CH242.15G	c 208	25	0.6	145	6	CD769674	AGENCOURT	CD769674
c 136	27	0.7	691	10	CM861380	8MS2kd48-	c 209	25	0.6	146	5	B0534815	AGENCOURT	B0534815
c 137	26	0.7	193	10	CE781142	tigr-g88-	c 210	25	0.6	146	5	B0842656	AGENCOURT	B0842656
c 138	26	0.7	202	3	B1298055	UI-R-CV2-	c 211	25	0.6	146	6	CF511584	CABud0002	CF511584
c 139	26	0.7	210	6	CB857501	NISC.na07	c 212	25	0.6	147	5	B0530443	AGENCOURT	B0530443
c 140	26	0.7	210	6	CB857502	NISC.na07	c 213	25	0.6	147	6	CD768528	AGENCOURT	CD768528
c 141	26	0.7	220	1	AJ703441	AJ703441	c 214	25	0.6	148	5	B0844094	AGENCOURT	B0844094
c 142	26	0.7	321	3	BM031265	496621.MA	c 215	25	0.6	148	6	CB237919	AGENCOURT	CB237919
c 143	26	0.7	331	2	BE852434	u20d04.x	c 216	25	0.6	149	5	B0563335	AGENCOURT	B0563335
c 144	26	0.7	370	3	BM958744	PLATE.16	c 217	25	0.6	149	6	CD769174	AGENCOURT	CD769174
c 145	26	0.7	377	5	B0727434	UI-R-C01-	c 218	25	0.6	150	5	B0530986	AGENCOURT	B0530986
c 146	26	0.7	410	2	B1291331	UI-R-CV2-	c 219	25	0.6	150	5	B0531230	AGENCOURT	B0531230
c 147	26	0.7	424	10	CE447107	tigr-g88-	c 220	25	0.6	150	5	B0531608	AGENCOURT	B0531608
c 148	26	0.7	424	11	CR894413	Sub scrof	c 221	25	0.6	151	6	CB976825	CAB40003	CB976825
c 149	26	0.7	457	10	CE056288	tigr-g88-	c 222	25	0.6	151	6	CD769343	AGENCOURT	CD769343
c 150	26	0.7	463	7	CO609245	DG8-80e18	c 223	25	0.6	151	6	CD769494	AGENCOURT	CD769494
c 151	26	0.7	476	7	CV554309	UI-D-GC1-	c 224	25	0.6	152	5	B0531635	AGENCOURT	B0531635
c 152	26	0.7	512	5	BX529297	BX529297	c 225	25	0.6	152	6	CD768990	AGENCOURT	CD768990
c 153	26	0.7	515	11	CR922091	Sub scrof	c 226	25	0.6	153	5	B0529255	AGENCOURT	B0529255
c 154	26	0.7	540	9	AZ869000	2M0180M20	c 227	25	0.6	153	5	B0554957	AGENCOURT	B0554957
c 155	26	0.7	543	5	B0680299	UI-CF-DU1	c 228	25	0.6	154	5	B0563541	AGENCOURT	B0563541
c 156	26	0.7	566	7	CO089397	GR_Ea08N	c 229	25	0.6	154	6	CD768221	AGENCOURT	CD768221
c 157	26	0.7	574	9	AZ554775	RPCT-23-2	c 230	25	0.6	154	6	CD769174	AGENCOURT	CD769174
c 158	26	0.7	596	11	CR894944	Sub scrof	c 231	25	0.6	155	6	CF512543	CABud0003	CF512543
c 159	26	0.7	597	1	AM637729	b161h07.w	c 232	25	0.6	156	1	AI345315	tb81e08.x	AI345315
c 160	26	0.7	597	7	CV016456	CV016456	c 233	25	0.6	159	5	B0556412	AGENCOURT	B0556412
c 161	26	0.7	599	7	CV585887	L AH-aaa3	c 234	25	0.6	159	6	CD769459	AGENCOURT	CD769459
c 162	26	0.7	600	7	CV585764	CV585764	c 235	25	0.6	162	5	B0556005	AGENCOURT	B0556005
c 163	26	0.7	613	6	CB975827	CAB40001	c 236	25	0.6	163	3	BM121849	L0502D03-	BM121849
c 164	26	0.7	618	5	B0680866	UI-CF-EC1	c 237	25	0.6	163	7	CN385655	LE2TR04D0	CN385655
c 165	26	0.7	631	5	B0623815	UI-H-FG1-	c 238	25	0.6	164	6	CB971610	CAB10005	CB971610
c 166	26	0.7	640	7	CN412664	170005325	c 239	25	0.6	165	7	CN385177	LE2TR0201	CN385177
c 167	26	0.7	646	3	BM980620	UI-CF-BN1	c 240	25	0.6	167	5	B0842470	AGENCOURT	B0842470
c 168	26	0.7	653	6	CA424914	UI-H-FB1-	c 241	25	0.6	168	5	B0958197	AGENCOURT	B0958197



388	25	0.6	565	8	DR148466	461	25	0.6	754	7	CK784592	CK784592 UI-D-GC1-
389	25	0.6	566	5	BX687988	462	25	0.6	755	5	BX761543	BX761543 RPT1543
390	25	0.6	568	2	BP401678	463	25	0.6	755	9	BH107066	BH107066 BPCT-24-3
C 391	25	0.6	570	6	CF212602	464	25	0.6	755	10	AG484585	AG484585 Mus muscu
C 392	25	0.6	570	8	DR405936	465	25	0.6	766	5	BX761207	BX761207 BX761207
C 393	25	0.6	574	2	BF729281	466	25	0.6	778	5	BX684116	BX684116 BX684116
C 394	25	0.6	579	6	CD776467	C 467	25	0.6	780	7	CO108108	CO108108 GR EB003
C 395	25	0.6	580	7	KN726470	468	25	0.6	787	1	AU117063	AU117063 AU117063
C 396	25	0.6	589	11	CR886685	C 469	25	0.6	796	5	BU589598	BU589598 AGENCOURT
C 397	25	0.6	594	5	BU677874	470	25	0.6	796	8	CX825761	CX825761 JGI CRAK4
C 398	25	0.6	595	1	AV717635	471	25	0.6	798	11	SSC561049	AJ561049 Sub scrof
C 399	25	0.6	597	6	CF206271	C 472	25	0.6	812	5	BU842689	BU842689 AGENCOURT
C 400	25	0.6	603	10	CE827724	473	25	0.6	812	6	CB235818	CB235818 AGENCOURT
C 401	25	0.6	607	5	BW961259	474	25	0.6	814	10	AG827284	AG827284 Sub scrof
C 402	25	0.6	609	1	AJ811318	475	25	0.6	815	5	BX756706	BX756706 BX756706
C 403	25	0.6	614	6	CD488859	C 476	25	0.6	817	5	CB998656	CB998656 AGENCOURT
C 404	25	0.6	615	8	CV916072	C 477	25	0.6	842	5	BX763164	BX763164 BX763164
C 405	25	0.6	616	6	CF215370	478	25	0.6	848	10	AG827256	AG827256 Sub scrof
C 406	25	0.6	615	7	KN384794	479	25	0.6	854	7	CR430666	CR430666 CR430666
C 407	25	0.6	617	11	CR149657	480	25	0.6	856	10	DU039766	DU039766 24696 Tom
C 408	25	0.6	620	2	BB631852	481	25	0.6	859	2	BG243773	BG243773 602357086
C 409	25	0.6	627	1	AL884520	482	25	0.6	866	7	CO385194	CO385194 AGENCOURT
C 410	25	0.6	633	3	BQ206970	483	25	0.6	866	10	AG827188	AG827188 Sub scrof
C 411	25	0.6	634	7	CF320026	484	25	0.6	868	6	CD756153	CD756153 AGENCOURT
C 412	25	0.6	635	2	BF132027	C 485	25	0.6	872	7	CK194558	CK194558 FGAS00298
C 413	25	0.6	637	2	BF213454	C 486	25	0.6	879	5	BU588050	BU588050 AGENCOURT
C 414	25	0.6	640	7	KN433884	C 487	25	0.6	884	7	CK814312	CK814312 Raqsgc472
C 415	25	0.6	644	7	KN385045	C 488	25	0.6	884	7	KN384923	KN384923 LE2TR02D2
C 416	25	0.6	644	7	KN385216	489	25	0.6	886	7	CK157081	CK157081 FGAS03815
C 417	25	0.6	645	8	DR405234	490	25	0.6	886	8	DN583815	DN583815 90328688
C 418	25	0.6	649	7	CK450028	C 491	25	0.6	898	2	BE878913	BE878913 601492614
C 419	25	0.6	658	2	BB630813	C 492	25	0.6	904	7	KN384845	KN384845 LE2TR02A1
C 420	25	0.6	660	6	CB339464	493	25	0.6	907	6	CA789790	CA789790 AGENCOURT
C 421	25	0.6	661	8	DR472524	C 494	25	0.6	924	10	AJ885370	AJ885370 Equus cab
C 422	25	0.6	661	4	AY440238	C 495	25	0.6	924	10	CZ505378	CZ505378 GWM2-10J1
C 423	25	0.6	662	9	AQ488364	496	25	0.6	944	10	DU050209	DU050209 144777 To
C 424	25	0.6	663	6	CF794201	C 497	25	0.6	949	7	CO387678	CO387678 AGENCOURT
C 425	25	0.6	665	8	DN128468	C 498	25	0.6	1070	5	BU300629	BU300629 603611639
C 426	25	0.6	666	8	CX448696	C 499	25	0.6	1096	9	CC287888	CC287888 CH261-107
C 427	25	0.6	671	6	CF792835	C 500	25	0.6	1101	10	CNS000FXS	AL070916 Drosophill
C 428	25	0.6	674	1	AL901343	C 501	25	0.6	1121	9	CC201777	CC201777 CH261-5C4
C 429	25	0.6	680	6	CB983173	C 502	25	0.6	1148	1	AL567420	AL567420 AL567420
C 430	25	0.6	680	11	CR924367	C 503	25	0.6	1199	8	DR145973	DR145973 49098853
C 431	25	0.6	685	1	AV733869	C 504	25	0.6	1244	8	DN809815	DN809815 77080735
C 432	25	0.6	687	8	DR575043	C 505	25	0.6	1248	3	BM452402	BM452402 AGENCOURT
C 433	25	0.6	688	6	CB241449	C 506	25	0.6	1503	2	BG500956	BG500956 602546539
C 434	25	0.6	688	11	CR903632	C 507	25	0.6	1805	10	CL078548	CL078548 CH216-151
C 435	25	0.6	692	3	BM993076	C 508	25	0.6	2222	4	BC018960	BC018960 Homo sapi
C 436	25	0.6	692	7	CK814475	C 509	25	0.6	4021	4	CNS0A4LN	BF632484 NF039G06D
C 437	25	0.6	692	10	CL367328	510	24	0.6	52	2	BF632484	BF632484 NF039G06D
C 438	25	0.6	693	8	CR895119	C 511	24	0.6	69	7	CV623286	CV623286 L_AJ-aaa1
C 439	25	0.6	703	6	CD237523	512	24	0.6	70	7	CV623393	CV623393 L_AJ-aaa1
C 440	25	0.6	703	7	KN106864	C 513	24	0.6	109	7	CV615513	CV615513 L_AJ-aaa2
C 441	25	0.6	707	1	AV716035	C 514	24	0.6	116	6	CB344090	CB344090 CA48EN000
C 442	25	0.6	707	7	KN385296	C 515	24	0.6	121	1	AI1305645	AI1305645 Qw72909.x
C 443	25	0.6	708	5	BX691192	C 516	24	0.6	127	2	BI135249	BI135249 UI-M-BH3-
C 444	25	0.6	712	11	CR267126	517	24	0.6	134	6	CB344144	CB344144 CA48EN000
C 445	25	0.6	714	7	CK738577	C 518	24	0.6	138	1	AI224731	AI224731 QX25C09.x
C 446	25	0.6	714	11	CR922773	C 519	24	0.6	158	3	BP145280	BP145280 BP145280
C 447	25	0.6	721	6	CD238347	C 520	24	0.6	166	5	BY082908	BY082908 BY082908
C 448	25	0.6	724	6	CD238943	C 521	24	0.6	167	5	BY590216	BY590216 AGENCOURT
C 449	25	0.6	725	5	BU850746	522	24	0.6	168	2	BE110172	BE110172 UI-R-BJ1-
C 450	25	0.6	728	11	CR800132	C 523	24	0.6	178	2	BP215451	BP215451 601881004
C 451	25	0.6	729	6	CD239530	C 524	24	0.6	179	11	CR910000	CR910000 Sub scrof
C 452	25	0.6	730	7	KN047319	525	24	0.6	180	2	BF508454	BF508454 UI-H-B14-
C 453	25	0.6	732	6	CB982531	C 526	24	0.6	191	2	BF733479	BF733479 MRO-AN003
C 454	25	0.6	732	8	DR811302	C 527	24	0.6	199	1	AA910608	AA910608 OK61a06.s
C 455	25	0.6	736	7	CK811424	C 528	24	0.6	199	1	AI251235	AI251235 QV55G01.x
C 456	25	0.6	740	6	CD238545	C 529	24	0.6	199	1	AI733997	AI733997 QV55G01.x
C 457	25	0.6	740	10	CE553122	530	24	0.6	204	2	BF402224	BF402224 UI-R-CA1-
C 458	25	0.6	743	9	CC310782	C 531	24	0.6	204	10	CL346304	CL346304 RPTC14-27
C 459	25	0.6	744	7	CV276358	532	24	0.6	212	10	CL361144	CL361144 RPTC14-34
C 460	25	0.6	749	11	CR194187	533	24	0.6	215	7	CO220941	CO220941 WS01011.B



534	24	0.6	221	1	AW522595	UI-R-B00-	607	24	0.6	432	3	BQ032062	BQ032062	UI-1-CF0-
c 535	24	0.6	225	1	AI046577	uh27e06.r	c 608	24	0.6	433	1	AW457182	AW457182	UI-M-BH3-
536	24	0.6	228	1	AI792898	qV55g01.Y	609	24	0.6	434	1	AV740541	AV740541	AV740541
537	24	0.6	231	2	BE111168	UI-R-BU1-	610	24	0.6	439	1	AW632847	AW632847	BL101a07.X
c 538	24	0.6	232	8	DR904169	JGI XZT65	c 611	24	0.6	441	7	CN644374	CN644374	LIUUMIGEN
539	24	0.6	245	8	DN104664	1107462 M	c 612	24	0.6	442	2	BG348216	BG348216	dg59f08.Y
540	24	0.6	247	10	CL330693	CH242.4P1	c 613	24	0.6	443	11	CR902053	CR902053	Sus scroF
541	24	0.6	249	1	AL962690	AL962690	614	24	0.6	445	2	BE490944	BE490944	dB40b10.X
542	24	0.6	249	1	AW528512	UI-R-B01-	615	24	0.6	449	3	BQ031536	BQ031536	UI-1-CF0-
c 543	24	0.6	251	1	AA823304	VP37C07.r	616	24	0.6	450	1	AW298014	AW298014	UI-H-BW0-
544	24	0.6	256	7	CN267265	170004708	c 617	24	0.6	450	8	DR050882	DR050882	MBF0003
545	24	0.6	257	1	AI568036	tr88e10.X	618	24	0.6	452	1	AW074271	AW074271	xB70c10.X
546	24	0.6	259	8	DN110163	1107825 M	c 619	24	0.6	456	1	AJ700742	AJ700742	AJ700742
547	24	0.6	259	8	DN110511	1108209 M	c 620	24	0.6	456	8	DN139438	DN139438	SGF264276
c 548	24	0.6	263	2	BS562944	BS562944	621	24	0.6	458	1	AI885486	AI885486	WM24f05.X
549	24	0.6	265	5	BUT35805	UI-E-DX0-	622	24	0.6	463	1	AJ280852	AJ280852	4A3A-AY-
c 550	24	0.6	269	3	BM812123	fx14f06.Y	623	24	0.6	463	1	AW136703	AW136703	UI-H-B11-
c 551	24	0.6	272	2	BI044595	PMO-OT021	624	24	0.6	465	5	BQ391277	BQ391277	NISC mql7
552	24	0.6	274	1	AL875163	AL875163	625	24	0.6	466	9	AQ310697	AQ310697	CITB1-El-
553	24	0.6	282	1	AA516315	ng88h11.8	626	24	0.6	467	9	BH552392	BH552392	BOGNF66TR
554	24	0.6	290	1	AW292572	UI-H-B12-	627	24	0.6	469	5	BY566085	BY566085	BY566085
c 555	24	0.6	291	7	CN083700	EC2BBA23A	628	24	0.6	475	5	EX107687	EX107687	EX107687
c 556	24	0.6	294	1	AA355995	EST64648	629	24	0.6	475	5	EX568221	EX568221	BX568221
c 557	24	0.6	296	6	CF722804	E04 LKFP00	630	24	0.6	475	10	CW066231	CW066231	104 313 1
558	24	0.6	299	10	BX289619	BX289619	631	24	0.6	476	7	CK784232	CK784232	UI-D-GC1-
c 559	24	0.6	302	8	DN104986	1100846 M	632	24	0.6	482	10	CL370447	CL370447	RPC144.29
560	24	0.6	305	3	BI296776	UI-R-DK0-	633	24	0.6	491	5	EX762998	EX762998	BX762998
561	24	0.6	306	1	AI355451	qt95e07.X	c 634	24	0.6	497	3	BM714864	BM714864	UI-E-E70-
562	24	0.6	310	1	AI955112	wt48h04.X	c 635	24	0.6	497	3	BM935083	BM935083	UI-M-BH3-
563	24	0.6	315	10	CL411220	RPC144.42	c 636	24	0.6	499	3	BM675378	BM675378	UI-E-E70-
c 564	24	0.6	318	10	CL370729	RPC144.29	c 637	24	0.6	499	9	BH090249	BH090249	RPCI-24-3
565	24	0.6	321	10	AG203964	Pan trogl	638	24	0.6	500	9	CE151011	CE151011	tigr-g88-
c 566	24	0.6	323	1	AI593222	vt86b02.X	639	24	0.6	505	1	AW632884	AW632884	bl01d08.X
567	24	0.6	324	1	AW321058	uo17e11.X	640	24	0.6	506	1	AL134552	AL134552	DKEP547A
c 568	24	0.6	327	10	CE797793	tigr-g88-	641	24	0.6	506	8	DR103310	DR103310	JHU030A10
569	24	0.6	336	1	AW449516	UI-H-B13-	642	24	0.6	512	5	BX113834	BX113834	BX113834
c 570	24	0.6	337	2	BP661503	ma87805.	643	24	0.6	512	10	CW490786	CW490786	fsb001f2
c 571	24	0.6	339	6	CB084012	hq07h06.B	c 644	24	0.6	516	9	AQ207299	AQ207299	HS_3242-A
572	24	0.6	342	8	CX125445	1326056 N	c 645	24	0.6	519	3	BI379779	BI379779	BF1G1_001
573	24	0.6	348	1	AW016332	UI-H-B10-	c 646	24	0.6	520	5	BUR24240	BUR24240	UBG2BPA07
574	24	0.6	348	10	CL362961	CL362961	c 647	24	0.6	521	8	DR130839	DR130839	49078376
575	24	0.6	362	1	AW048768	UI-M-BH1-	648	24	0.6	526	1	AA997646	AA997646	UI-R-CO-h
c 576	24	0.6	363	10	CE411596	tigr-g88-	649	24	0.6	529	5	BUR24195	BUR24195	UBG1BFP06
c 577	24	0.6	364	2	BE234499	141628 MA	650	24	0.6	532	2	BG379298	BG379298	UI-R-CS0-
578	24	0.6	365	3	BQ160489	WHE1821-1	c 651	24	0.6	532	8	CX693530	CX693530	YDC82h04.
579	24	0.6	367	8	DN139422	SGP264260	c 652	24	0.6	533	1	AM004437	AM004437	AM004437
c 580	24	0.6	373	3	BP804554	BP804554	653	24	0.6	533	11	DR19GAT	DR19GAT	Danio rer
581	24	0.6	374	6	CA296841	SCUTS208	654	24	0.6	536	8	DN435011	DN435011	LIB84217-1
582	24	0.6	381	1	AI044248	UI-R-Cl-k	c 655	24	0.6	536	10	AG829138	AG829138	Sus scroF
583	24	0.6	382	1	AI706215	UI-R-AB1-	656	24	0.6	537	2	BF389909	BF389909	UI-R-B82-
584	24	0.6	382	1	AI161157	QC89h04.X	657	24	0.6	537	7	CO000974	CO000974	OB073 pre
585	24	0.6	386	5	BY097208	BY097208	c 658	24	0.6	537	8	DR033904	DR033904	OD2 pre-o
586	24	0.6	388	10	CL352603	RPC144.40	c 659	24	0.6	539	5	BQ398314	BQ398314	NISC-mo06
c 587	24	0.6	390	11	CR899706	Sus scroF	660	24	0.6	544	9	CE197627	CE197627	tigr-g88-
588	24	0.6	396	5	BQ601516	MI-P-H4-a	c 661	24	0.6	545	10	CZ563848	CZ563848	1_P-BW039
589	24	0.6	399	5	BUR22195	UBG22195	662	24	0.6	546	11	CR872468	CR872468	Sus scroF
590	24	0.6	401	3	BI403623	MI-P-CP1-	c 663	24	0.6	547	1	AL036011	AL036011	DKEP564L
591	24	0.6	402	1	AI619980	ty46g01.X	664	24	0.6	548	2	BF704508	BF704508	MI-P-H4-a
592	24	0.6	404	1	AI521110	t043b08.X	665	24	0.6	550	9	AQ930618	AQ930618	RPCI-23-2
c 593	24	0.6	405	3	BI388085	BI388085	666	24	0.6	552	9	CE251519	CE251519	tigr-g88-
594	24	0.6	405	11	CR872320	Sus scroF	667	24	0.6	554	10	CL345058	CL345058	RPCI44.27
595	24	0.6	406	1	AI032161	OE76f08.8	c 668	24	0.6	555	9	AZ912824	AZ912824	RPCI-24-1
596	24	0.6	406	5	BX484831	DKEP2686K	c 669	24	0.6	556	9	CE135408	CE135408	tigr-g88-
597	24	0.6	407	3	BM151093	TCBAP1D13	c 670	24	0.6	558	8	DR134364	DR134364	49299044
c 598	24	0.6	411	3	BP650014	BP650014	671	24	0.6	559	8	DR142805	DR142805	49101185
599	24	0.6	413	1	AV744909	AV744909	672	24	0.6	559	8	DR151379	DR151379	49336343
600	24	0.6	416	1	AW055257	wz16a09.X	673	24	0.6	560	2	BF711910	BF711910	MI-P-O3-a
601	24	0.6	416	2	BE114328	UI-R-BU1-	674	24	0.6	563	7	CK433846	CK433846	UI-D-GC1-
602	24	0.6	418	1	AI043943	UI-R-CO-j	c 675	24	0.6	563	8	DR137230	DR137230	49296983
603	24	0.6	420	6	CB241920	UI-CF-FN0	676	24	0.6	563	11	CR873261	CR873261	Sus scroF
604	24	0.6	424	2	BF199042	249086 MA	677	24	0.6	565	3	BM987543	BM987543	UI-H-CO0-
605	24	0.6	425	7	CK785785	UI-D-GC1-	c 678	24	0.6	565	8	DR138826	DR138826	49159096
606	24	0.6	432	1	AW534664	UI-R-B80-	679	24	0.6	565	11	CR865597	CR865597	Sus scroF

CX796393	UI-R-DQ1-
CX138227	1277433 N
CR866681	Sus scrofa
CR914697	Sus scrofa
BB865302	BB865302
BE238970	MD0626 Me
CR420555	CR420555
CA811313	CA221011
CX716569	1332996 N
BB865486	BB865486
BPI159037	BPI159037
BZ026169	oei54a06.
BZ026169	Forward s
CX146993	1286945 N
BL1382917	BFLG2_001
CX719257	1335928 N
CX538188	CH240_417
CX375037	JGI_XZT58
AG184527	Pan trogl
AG301460	Mus muscu
CX497433	JGI_XZG43
CR050194	Reverse s
CX363975	638790 NC
BU1002X1	BU1002X1
WS0201.B	WS0201.B
CV269295	WS0202.B2
CV269520	WS0208.B2
CR415639	CR415639
CR925212	Sus scrofa
WS02029.B	WS02029.B
CR423544	CR423544
BP455435	BP455435
CN972580	20077_124
RPC1-24-3	RPC1-24-3
BH086892	JGI_CABAL
DNO57376	JGI_CABAL
DNO57377	JGI_CABAL
DE099057	Oryzizias 1
UI-H-DH1-	UI-H-DH1-
BW976827	BW976827
BM76827	BM76827
1335428 N	1335428 N
CX590599	CH240_390
UI-R-F80-	UI-R-F80-
BM762190	BM762190
UI-R-E00-	UI-R-E00-
CR906699	Sus scrofa
CR437457	CR437457
BW969191	BW969191
CD000612	AGENCOURT
CR425016	CR425016
Single re	Single re
JGI_XZG62	JGI_XZG62
BX781246	BX781246
BX876111	JGI_CAA11
BX734302	BX734302
UI-M-HK0-	UI-M-HK0-
BZ160889	CH230-230
AJ724457	AJ724457
nad22d02.	nad22d02.
AG828972	Sus scrofa
CR585029	CR585029
BQ771803	UI-H-E21-
AD261886	Tetraodon
CD513863	AGENCOURT
BU559835	AGENCOURT
DN113028	1112267 M
AWB74684	kdef2 Soa
CX625298	H29_F3 Sp
BX742745	BX742745
CO913873	AGENCOURT
AZ688419	ENTL116TR
CX370217	JGI_XZT55
CC535951	CH240_460
CR184556	Forward

826	24	0.6	877	11	CR882784	Sus scrofa	CR882784 Sus scrofa	C 899	23	0.6	198	7	CF888246
827	24	0.6	885	7	CK019779	AGENCOURT	CK019779 AGENCOURT	C 900	23	0.6	201	1	AI933727
828	24	0.6	887	2	BI101506	602887305	BI101506 602887305	C 901	23	0.6	204	2	BF577654
829	24	0.6	889	5	BU854717	AGENCOURT	BU854717 AGENCOURT	C 902	23	0.6	213	2	EG296420
830	24	0.6	900	5	BU164249	AGENCOURT	BU164249 AGENCOURT	C 903	23	0.6	214	5	BX734913
831	24	0.6	920	11	SSCS14471		AJ514471 Sus scrofa	C 904	23	0.6	218	1	AI501906
832	24	0.6	938	5	BQ880486	AGENCOURT	BQ880486 AGENCOURT	C 905	23	0.6	220	1	AW531369
833	24	0.6	952	1	AF179479	AGENCOURT	AF179479 AF179479	C 906	23	0.6	221	2	BF709767
834	24	0.6	962	8	DN562478	88066682	DN562478 88066682	C 907	23	0.6	222	1	AW413699
835	24	0.6	983	2	BE957745	601653858	BE957745 601653858	C 908	23	0.6	222	2	BF933259
836	24	0.6	1027	5	BQ939830	AGENCOURT	BQ939830 AGENCOURT	C 909	23	0.6	225	5	BU198036
837	24	0.6	1052	8	DR134327	49245176	DR134327 49245176	C 910	23	0.6	225	5	BU597618
838	24	0.6	1099	8	CR817722	D12 AGCS	CR817722 D12 AGCS	C 911	23	0.6	230	6	CD572080
839	24	0.6	1107	8	DR740855	FGS00078	DR740855 FGS00078	C 912	23	0.6	231	1	AW049251
840	24	0.6	1153	8	DR123662	49175497	DR123662 49175497	C 913	23	0.6	232	3	BQ023848
841	24	0.6	1166	7	CV672916	RET7SJ 09	CV672916 RET7SJ 09	C 914	23	0.6	240	1	AI418803
842	24	0.6	1172	8	DN262630	Meso03860	DN262630 Meso03860	C 915	23	0.6	241	2	BG447004
843	24	0.6	1201	10	CNS027L1		AL184798 Tetraodon	C 916	23	0.6	242	1	AI844285
844	24	0.6	1249	8	CR100315	RECm1168	CR100315 RECm1168	C 917	23	0.6	245	2	BB579813
845	24	0.6	1307	10	AG346257	Mus muscu	AG346257 Mus muscu	C 918	23	0.6	249	1	AI071623
846	24	0.6	1473	10	AJ860081	Brassica	AJ860081 Brassica	C 919	23	0.6	250	2	BE119256
847	24	0.6	1739	10	CL136004	ISB1-107F	CL136004 ISB1-107F	C 920	23	0.6	252	2	BF392547
848	24	0.6	1961	4	AK051402	Mus muscu	AK051402 Mus muscu	C 921	23	0.6	253	7	CK468782
849	24	0.6	2838	4	CR858400	Pongo pyg	CR858400 Pongo pyg	C 922	23	0.6	254	1	AV766309
850	24	0.6	2956	4	HSM806920	Homo sapi.	EX640812 Homo sapi.	C 923	23	0.6	254	5	BU728008
851	23	0.6	59	6	CB217739	NISC nb04	CB217739 NISC nb04	C 924	23	0.6	257	1	AW121760
852	23	0.6	79	1	AJ737158	AJ737158	AJ737158 AJ737158	C 925	23	0.6	257	5	BU555340
853	23	0.6	97	5	BQ526869	NISC no19	BQ526869 NISC no19	C 926	23	0.6	258	1	AA877929
854	23	0.6	103	7	CN545910	EST 17854	CN545910 EST 17854	C 927	23	0.6	259	1	AJ759046
855	23	0.6	105	1	AM006188	AM006188	AM006188 AM006188	C 928	23	0.6	260	1	AV736595
856	23	0.6	111	6	CA811895	CA811895	CA811895 CA811895	C 929	23	0.6	261	8	CK764081
857	23	0.6	120	6	CF333073	JMT--01-N	CF333073 JMT--01-N	C 930	23	0.6	263	10	CL345026
858	23	0.6	121	8	CK769791	ACAD-aab1	CK769791 ACAD-aab1	C 931	23	0.6	265	3	BI398887
859	23	0.6	124	1	AI540464	tg24c10.x	AI540464 tg24c10.x	C 932	23	0.6	265	10	CL323384
860	23	0.6	125	6	CA812378	CA41LN031	CA812378 CA41LN031	C 933	23	0.6	266	1	AW125065
861	23	0.6	127	11	CR897891	Sus scrofa	CR897891 Sus scrofa	C 934	23	0.6	266	7	CK784345
862	23	0.6	129	3	AZ077736	RPCI-23-4	AZ077736 RPCI-23-4	C 935	23	0.6	266	10	CL412219
863	23	0.6	136	1	AI540534	tg25f05.x	AI540534 tg25f05.x	C 936	23	0.6	268	7	CO535869
864	23	0.6	145	1	AI843728	UI-R-AK1	AI843728 UI-R-AK1	C 937	23	0.6	273	8	CK067523
865	23	0.6	149	7	CV600751	L AH-aaal	CV600751 L AH-aaal	C 938	23	0.6	275	7	CJ025364
866	23	0.6	150	7	CV600885	L AH-aaal	CV600885 L AH-aaal	C 939	23	0.6	275	10	CE311418
867	23	0.6	150	8	DN586294	EST6285 Z	DN586294 EST6285 Z	C 940	23	0.6	277	1	AW014864
868	23	0.6	153	1	AA040237	ZK44a10.s	AA040237 ZK44a10.s	C 941	23	0.6	278	5	BU754788
869	23	0.6	157	5	BN957428	BN957428	BN957428 BN957428	C 942	23	0.6	279	1	AI709687
870	23	0.6	158	1	AI058697	UI-R-CL-k	AI058697 UI-R-CL-k	C 943	23	0.6	283	10	CL409223
871	23	0.6	158	9	AZ723254	RPCI-24-6	AZ723254 RPCI-24-6	C 944	23	0.6	286	3	BI302910
872	23	0.6	159	2	BG957829	PM2-CT080	BG957829 PM2-CT080	C 945	23	0.6	288	6	CB710753
873	23	0.6	166	7	CK362261	AGENCOURT	CK362261 AGENCOURT	C 946	23	0.6	288	7	CK349134
874	23	0.6	168	9	AZ819262	2M0089J14	AZ819262 2M0089J14	C 947	23	0.6	290	5	BN955926
875	23	0.6	171	3	BP460907	BP460907	BP460907 BP460907	C 948	23	0.6	290	10	CE745191
876	23	0.6	172	1	AI560685	tg55b01.x	AI560685 tg55b01.x	C 949	23	0.6	292	1	AW277138
877	23	0.6	178	3	BP154966	BP154966	BP154966 BP154966	C 950	23	0.6	295	3	BI320169
878	23	0.6	180	1	AL597503	DKF2p313E	AL597503 DKF2p313E	C 951	23	0.6	295	10	CL354606
879	23	0.6	181	5	BU860664	AGENCOURT	BU860664 AGENCOURT	C 952	23	0.6	296	2	BB477231
880	23	0.6	182	8	CK907949	JGI CAAN6	CK907949 JGI CAAN6	C 953	23	0.6	297	1	AI202045
881	23	0.6	183	5	BU540127	AGENCOURT	BU540127 AGENCOURT	C 954	23	0.6	297	3	BI859231
882	23	0.6	185	3	BQ276832	AGENCOURT	BQ276832 AGENCOURT	C 955	23	0.6	297	7	CF959179
883	23	0.6	185	5	BQ596755	PESToab2	BQ596755 PESToab2	C 956	23	0.6	298	1	AW078853
884	23	0.6	185	5	BU565876	AGENCOURT	BU565876 AGENCOURT	C 957	23	0.6	298	5	BU755026
885	23	0.6	185	5	BU587367	AGENCOURT	BU587367 AGENCOURT	C 958	23	0.6	298	10	AL760959
886	23	0.6	185	5	BU928581	AGENCOURT	BU928581 AGENCOURT	C 959	23	0.6	299	3	BP436094
887	23	0.6	185	5	BU942925	AGENCOURT	BU942925 AGENCOURT	C 960	23	0.6	301	1	AI555577
888	23	0.6	186	5	BU852006	AGENCOURT	BU852006 AGENCOURT	C 961	23	0.6	302	1	AI348747
889	23	0.6	186	5	BU928649	AGENCOURT	BU928649 AGENCOURT	C 962	23	0.6	303	1	AI043915
890	23	0.6	186	11	CR887441	Sus scrofa	CR887441 Sus scrofa	C 963	23	0.6	303	2	BI105607
891	23	0.6	189	1	AA746619	rx27a05.s	AA746619 rx27a05.s	C 964	23	0.6	306	1	AI223445
892	23	0.6	192	5	BU536052	AGENCOURT	BU536052 AGENCOURT	C 965	23	0.6	307	3	BP156398
893	23	0.6	193	5	BU960980	AGENCOURT	BU960980 AGENCOURT	C 966	23	0.6	307	8	CV942003
894	23	0.6	195	2	BE804639	8x84e07.y	BE804639 8x84e07.y	C 967	23	0.6	309	10	CE534015
895	23	0.6	195	5	BU564300	AGENCOURT	BU564300 AGENCOURT	C 968	23	0.6	311	8	CK067671
896	23	0.6	196	5	BU928948	AGENCOURT	BU928948 AGENCOURT	C 969	23	0.6	312	3	BP157861
897	23	0.6	197	5	BU554869	AGENCOURT	BU554869 AGENCOURT	C 970	23	0.6	313	3	BM987127
898	23	0.6	198	3	BI298639	UI-R-CV2	BI298639 UI-R-CV2	C 971	23	0.6	314	1	AW792746



Db 1957 CAAGGCACAGGAGGAAAGTCAATTCAGACTGATGAGTGGAGAAATGCCCTGTGCGAA 2016  
Qy 2080 GAACGCTTGAGTATGCCCTTGGAAGGGCATTGAAAAACATATATTATGAGGATCTGAG 2139  
Db 2017 GAACGCTTGAGTATGCCCTTGGAAGGGCATTGAAAAACATATATTATGAGGATCTGAG 2076  
Qy 2140 GAAGCCAGGTTAAACCAAAAAAATATCCCGACCTCTCAATATATTAAGGACCCCTG 2199  
Db 2077 GAAGCCAGGTTAAACCAAAAAAATATCCCGACCTCTCAATATATTAAGGACCCCTG 2136  
Qy 2200 ATGAATGGAATGAAATTTGTTGGTATCTTTTGGAGCTGGAATAATTTCTACCTCAG 2259  
Db 2137 ATGATGGAATGAAATTTGTTGGTATCTTTTGGAGCTGGAATAATTTCTACCTCAG 2196  
Qy 2260 GTTATAAAGTCAGCCCGGTTATGAAGAAGCTGTGGCCACCTTATCCCTTTCAATGAA 2319  
Db 2197 GTTATAAAGTCAGCCCGGTTATGAAGAAGCTGTGGCCACCTTATCCCTTTCAATGAA 2256  
Qy 2320 AAAGAAGAGAGAAACAGAGTCTTAAAGGACAGTACAGAGAGGAGGACCCCTTACCAG 2379  
Db 2257 AAAGAAGAGAGAAACAGAGTCTTAAAGGACAGTACAGAGAGGAGGACCCCTTACCAG 2316  
Qy 2380 GGACACCTGCTGCTGGCCACTGTTAAAGGCGACGTGACGACATAGGCAAGAAACATAGTT 2439  
Db 2317 GGACACCTGCTGCTGGCCACTGTTAAAGGCGACGTGACGACATAGGCAAGAAACATAGTT 2376  
Qy 2440 GGAGTAGCTCTGGCTGCAATAATTTCCGAGTTATTTAGTATGAGGATCATGACTCCATGT 2499  
Db 2377 GGAGTAGCTCTGGCTGCAATAATTTCCGAGTTATTTAGTATGAGGATCATGACTCCATGT 2436  
Qy 2500 GATGAAGATCTGAAGCTGCTTGACCAACAGCAGATATATTTGSCCTGTACGACTC 2559  
Db 2437 GATAAGATCTGAAGCTGCTTGACCAACAGCAGATATATTTGSCCTGTACGACTC 2496  
Qy 2560 ATCACTCTCTCCCTGGATGAAATGATTTTGTCCAGGAAATGGAGAGATTAGCTATA 2619  
Db 2497 ATCACTCTCTCCCTGGATGAAATGATTTTGTCCAGGAAATGGAGAGATTAGCTATA 2556  
Qy 2620 AGGATTCATTTGTTGATGGAGAGCAACCACTTCAAAAAACCCACACAGCAGTTTAAATA 2679  
Db 2557 AGGATTCATTTGTTGATGGAGAGCAACCACTTCAAAAAACCCACACAGCAGTTTAAATA 2616  
Qy 2680 GCTCCGAGATACAGTGCACTGTAATCATGTCTCGACGCGTCCAAAGAGTGTGGTGGT 2739  
Db 2617 GCTCCGAGATACAGTGCACTGTAATCATGTCTCGACGCGTCCAAAGAGTGTGGTGGT 2676  
Qy 2740 TGTTCAGCTGTTAGATGAAATCTAAAGATGAATTAATTTGAGGAAATCATCGAAGAA 2799  
Db 2677 TGTTCAGCTGTTAGATGAAATCTAAAGATGAATTAATTTGAGGAAATCATCGAAGAA 2736  
Qy 2800 TATGAAGATATTAGACAGGACCATTTAGTGTCTCTCAAGGAGAGGAGATCTTACCCTTA 2859  
Db 2737 TATGAAGATATTAGACAGGACCATTTAGTGTCTCTCAAGGAGAGGAGATCTTACCCTTA 2796  
Qy 2860 AGTCAAGCCAGAAAGTGGTTTCAATGATGTTGCTGTGAACTCACCCTCAGGTAAG 2919  
Db 2797 AGTCAAGCCAGAAAGTGGTTTCAATGATGTTGCTGTGAACTCACCCTCAGGTAAG 2856  
Qy 2920 CCCACGTTTATGGACCCAGCTCTTGAAGACTATGACCTGCGAGAGCTGTTGACTAC 2979  
Db 2857 CCCACGTTTATGGACCCAGCTCTTGAAGACTATGACCTGCGAGAGCTGTTGACTAC 2916  
Qy 2980 ATTCACTGGAAGCTTTCTTTGATGCTGCGAGCTCCGGGCAAGTACCCGAATCGAGGC 3039  
Db 2917 ATTCACTGGAAGCTTTCTTTGATGCTGCGAGCTCCGGGCAAGTACCCGAATCGAGGC 2976  
Qy 3040 TTCCCCAAGATATTAAACGACAAAACAGTAGTGGAGGCGCAGGAAAGTCTACATGAT 3099  
Db 2977 TTCCCCAAGATATTAAACGACAAAACAGTAGTGGAGGCGCAGGAAAGTCTACATGAT 3036  
Qy 3100 GCCACATATGCTGAACACACTGATTAAGTCAAAAGAAATCCCGGGCCGGGTGTGGTT 3159  
Db 3037 GCCACATATGCTGAACACACTGATTAAGTCAAAAGAAATCCCGGGCCGGGTGTGGTT 3096

Qy 3160 GGTTCTGGCCAGCACAGAGATATCCAAAGACGACATTCACCTGTACGAGAGGCTGCTGTG 3219  
Db 3097 GGTTCTGGCCAGCACAGAGATATCCAAAGACGACATTCACCTGTACGAGAGGCTGCTGTG 3156  
Qy 3220 CCCACGCTGCAGAGCCCATAGCCACTTTCTATGGGTTTAAGGCAACAGGCTGAGAAGGAC 3279  
Db 3157 CCCACGCTGCAGAGCCCATAGCCACTTTCTATGGGTTTAAGGCAACAGGCTGAGAAGGAC 3216  
Qy 3280 TCTGCCAGCAGCGGAGCCATACCTACTGCTCTCAGACTTTCATCGCTCCCTTGCATTCGGC 3339  
Db 3217 TCTGCCAGCAGCGGAGCCATACCTACTGCTCTCAGACTTTCATCGCTCCCTTGCATTCGGC 3276  
Qy 3340 ATCCGTGACTACCTGGGCCCTGTTTGGCGTTCCTGCTTTGGGGTAGAAGAGCTGAGCAAG 3399  
Db 3277 ATCCGTGACTACCTGGGCCCTGTTTGGCGTTCCTGCTTTGGGGTAGAAGAGCTGAGCAAG 3336  
Qy 3400 GCTTATGAGGATGATGTCAGCACTACAGCAGCATCATGTCAAGGCGCTGGGGACCGG 3459  
Db 3337 GCTTATGAGGATGATGTCAGCACTACAGCAGCATCATGTCAAGGCGCTGGGGACCGG 3396  
Qy 3460 CTGGCAGAGGCTTTTCAGAGAAGCTCCCATGAAAGAGTTTCGCCGAGAAGCTGTGGGCTAC 3519  
Db 3397 CTGGCAGAGGCTTTTCAGAGAAGCTCCCATGAAAGAGTTTCGCCGAGAAGCTGTGGGCTAC 3456  
Qy 3520 TGTGGCAGTAGCAGCTGAGCTGCGAGACTCTCGAAGTGTGGGTACAAGGGGATCCGC 3579  
Db 3457 TGTGGCAGTAGCAGCTGAGCTGCGAGACTCTCGAAGTGTGGGTACAAGGGGATCCGC 3516  
Qy 3580 CCGGCTCTGGCTACCCAGCAGCCGACACACCGAGAAGCTCACCATGTGGAGACTC 3639  
Db 3517 CCGGCTCTGGCTACCCAGCAGCCGACACACCGAGAAGCTCACCATGTGGAGACTC 3576  
Qy 3640 GCAGACATCGAGCAGCTTACAGGCACTTAGGTAAACAGAAATCATTTAGCAATGGCACCTGCT 3699  
Db 3577 GCAGACATCGAGCAGCTTACAGGCACTTAGGTAAACAGAAATCATTTAGCAATGGCACCTGCT 3636  
Qy 3700 TCAGCAGCTTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAG 3759  
Db 3637 TCAGCAGCTTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAG 3696  
Qy 3760 ATTTCCAGGATCAG 3774  
Db 3697 ATTTCCAGGATCAG 3711

## RESULT 2

CR859882 LOCUS 4888 bp mRNA linear HTC 12-NOV-2004

DEFINITION Pongo pygmaeus mRNA; cDNA DKFp469P1116 (from clone DKFp469P1116).

ACCESSION CR859882

VERSION CR859882.1 GI:55730633

KEYWORDS HTC.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM

Pongo pygmaeus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Pongo.

1 (bases 1 to 4888)

OTterwaelde, B.; Obermaier, B.; Deutschbauer, S.; Schapp, A.,

Mewes, H.W., Weil, B., Amid, C., Oberger, A., Fobo, G., Han, M. and

Wiemann, S.

The German cDNA Consortium

Direct Submission

Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764

Neuherberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Medigenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project.

This clone (DKFp469P1116) is available at the RZPD Deutsches

Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.

Please contact RZPD for ordering.

http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp469P1116  
Further information about the clone and the sequencing project is  
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES  
Source  
1. .4888  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp469P1116"  
/tissue\_type="kidney"  
/clone\_lib="469 (synonym: pkid1). Vector pSport1\_sfi; host  
DH10B; sites SfiI + SfiIb"  
/dev\_stage="adult"  
/note="similar to 5-methyltetrahydrofolate--homocysteine  
methyltransferase (Homo sapiens)"  
1. .4888  
/gene="DKFZp469P1116"  
595. .2754  
/gene="DKFZp469P1116"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAH92038.1"  
/db\_xref="GI:55730634"  
/translation="MEENHLYAIFNFIHATKVIKELTPGARISGSLNLSFSGMEAI  
REAMQVFLYHAIKSGMDGIVNAGNLVYDDIHKELLQLEDILWNKQPEATEKLLR  
YACQGTGGKVKIOTDEWRNGSVEERLEYALVKGIEKHIVEDTEAKLNQEKYPRPLN  
IIEPLMNGMKIVGDLFGAGKQFLPQVIKSARVKKXAVGHLLIPWEKERETRVLNGT  
VEEDPYGTTIVLATVGDVHDIGKIVGVVGNFRVNDLGVMTPCDKILKAALDH  
KADIGLSLTPSLDENLFIKEMERLAIKPLIGATTSKTHTAVKAPRYSAV  
IHLVDASKSVVQSLDENLDFEIEIMEEDYIDQKHVESIKERRYPLPSQARKS  
FNDMLSEBPVPKPTFTGTVFDYDLQKLVYDWKPFVDMQLRGKVPNRGFPKI  
ENDTVGSEAKVVDVDAQNMLNTLISOKLQARGVGVFWPAQSLODDIHLVAREGAVQ  
AAEPIATFYGLRQAEKDSASTEPFYCLSDFIAPLHSGIRDYGLFPAVCFVZELSK  
AYEDDGDYSIMKALGDLRAEPABELHVRVRELMAYCGSQQLDVADLRLRYKG  
IRPAGPYQPDHTEKLTMRLADIQSTGRILRTESLAWAPASVGLYFSLNLSKSTYP  
AVGKISKQVEDYALRKNIISVAIEVKMLGPLILGDTD"

ORIGIN

Query Match	21.4%; Score 839; DB 4; Length 4888;
Best Local Similarity	98.9%; Pred. No. 0;
Matches 1789; Conservative	0; Mismatches 19; Indels 0; Gaps 0;
Qy 1747	AAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATTAAGTGGAGGTCCTTCCAACTGTGCC 1806
Db 640	AAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATTAAGTGGAGGTCCTTCCAACTGTGCC 699
Qy 1807	TTCTCTTCCGAGGAATGGAAGCCATTCCGAGAAGCAATGCATGGGGTTTTCTTTACCAT 1866
Db 700	TTCTCTTCCGAGGAATGGAAGCCATTCCGAGAAGCAATGCAGGGGGTTTTCTTTTACCAT 759
Qy 1867	GCAATCAAGTCGGCATGGACATGAGATAGTGAATGCTGGAAACCTCCCTGTGTATGAT 1926
Db 760	GCAATCAAGTCGGCATGGACATGAGATAGTGAATGCTGGAAACCTCCCTGTGTATGAT 819
Qy 1927	GATATCCATAGGAACCTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAGAACCCCTGAG 1986
Db 820	GATATCCATAGGAACCTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAGAACCCCTGAG 879
Qy 1987	GCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAGTCATT 2046
Db 880	GCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAGTCATT 939
Qy 2047	CAGACTGATGAGTGAAGAAATGGCCCTGTCGAGAACCGCTTGAGTATGCCCTTGTGAAG 2106
Db 940	CAGACTGATGAGTGAAGAAATGGCTCTGTGCGAAGAACGCTTGAGTATGCCCTTGTGAAG 999
Qy 2107	GGCACTTGAAAAACATATTATTGAGGATCTAGGAGGACAGGTATAAACCAAAAAAATAT 2166
Db 1000	GGCACTTGAAAAACATATTATTGAGGATCTAGGAGGACAGGTATAAACCAAAAAAATAT 1059
Qy 2167	CCCCGACCTCTCAATATATTAATGAAGACCCCTGTGATGAATGGAAATGAAATTTGTGTGAT 2226
Db 1060	CCCCGACCTCTCAATATATTAATGAAGACCCCTGTGATGAATGGAAATGAAATTTGTGTGAT 1119

Qy 2227	CTTTTTCGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAAGTCAGCCCGGGTTATGAAG 2286
Db 1120	CTTTTTCGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAAGTCAGCCCGGGTTATGAAG 1179
Qy 2287	AAGCTCTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAAACACAGAGTGCTT 2346
Db 1180	AAGCTCTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAAACACAGAGTGCTT 1239
Qy 2347	AACGGCACATAGNAGNAGAGGACCCCTTACACGGGCACCATCTGCTGGCCACCTGTTAAA 2406
Db 1240	AACGGCACATAGNAGNAGAGGACCCCTTACACGGGCACCATCTGCTGGCCACCTGTTAAA 1299
Qy 2407	GGCGACTGTCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCGCAATAATTC 2466
Db 1300	GGCGACTGTCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCGCAATAATTC 1359
Qy 2467	CGAGTTATTGATTTAGGAGTCATGACTCCATGTGATAAGATACCTGAAAGCTGCTTTGAC 2526
Db 1360	CGAGTTATTGATTTAGGAGTCATGACTCCATGTGATAAGATACCTGAAAGCTGCTTTGAC 1419
Qy 2527	CACAAACAGATATAATTGGCTGTGAGGACTCATCTCTTCCCTGGATGAATGATTT 2586
Db 1420	CACAAACAGATATAATTGGCTGTGAGGACTCATCTCTTCCCTGGATGAATGATTT 1479
Qy 2587	TTTGTTCGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGATTGGAGGAGCA 2646
Db 1480	TTTGTTCGCCAAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGATTGGAGGAGCA 1539
Qy 2647	ACCACTTCAAAAAACCCACACAGCAGATTAAATAGTCCGAGATACAGTGACCTGTAAATC 2706
Db 1540	ACCACTTCAAAAAACCCACACAGCAGATTAAATAGTCCGAGATACAGTGACCTGTAAATC 1599
Qy 2707	CATGCTCTGGACGGTCCAGAGTGCTGTGGTGTGTCCAGCTGTGTAGATGAATACTA 2766
Db 1600	CATGCTCTGGACGGTCCAGAGTGCTGTGGTGTGTCTCAGCTGTGTAGATGAATACTA 1659
Qy 2767	AAGCATGAATATTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAT 2826
Db 1660	AAGCATGAATATTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAT 1719
Qy 2827	GAGTCTCTCAAGGAGAGGAGATACCTTACCCCTTAAAGTCAAGCCAGAAAAAGTGGTTCCAA 2886
Db 1720	GAGTCTCTCAAGGAGAGGAGATACCTTACCCCTTAAAGTCAAGCCAGAAAAAGTGGTTCCAA 1779
Qy 2887	ATGATTTGGCTGTCTGAACTCACCGAGTGAAGCCACGTTTATTTGGACCCAGCTCTTT 2946
Db 1780	ATGATTTGGCTGTCTGAACTCACCGAGTGAAGCCACGTTTATTTGGACCCAGCTCTTT 1839
Qy 2947	GAAGACTATGACCTGCAGAGAGCTGCTGGACTACATTTGACTGGAAGCCCTTTCTTTGATGTC 3006
Db 1840	GAAGACTATGACCTGCAGAGAGCTGCTGGACTACATTTGACTGGAAGCCCTTTCTTTGATGTC 1899
Qy 3007	TGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCGCAAGATATTAAACGACAAAAACA 3066
Db 1900	TGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCGCAAGATATTAAACGACAAAAACA 1959
Qy 3067	GTAGGTGAGAGGCGCAGGAAGGTCACGATGATGCCCAATATGCTGAAACACACTGATTT 3126
Db 1960	GTAGGTGAGAGGCGCAGGAAGGTCACGATGATGCCCAATATGCTGAAACACACTGATTT 2019
Qy 3127	AGTCAAAAGAAATCCCGGGCCCGGGGTGTGGTTCGGTTCTGGCCAGCACAGAGTATCCAA 3186
Db 2020	AGTCAAAAGAAATCCCGGGCCCGGGGTGTGGTTCGGTTCTGGCCAGCACAGAGTATCCAA 2079
Qy 3187	GACGACATTCACCTGTACGACAGGCTGCTGTGCCCGCCAGGCTGCAGAGCCCATAGCCACT 3246
Db 2080	GATGACATTCACCTGTACGACAGGCTGCTGTGCCCGCCAGGCTGCAGAGCCCATAGCCACC 2139
Qy 3247	TTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCGCAGACGGAGCCCATACTACTGC 3306
Db 2140	TTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCGCAGACGGAGCCCATACTACTGC 2199





This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.

This clone (DKFZp313F151) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers  
1..721  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp313F151"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

## ORIGIN

Query Match 16.8%; Score 660; DB 1; Length 721;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1893 GATAGTGAATCTCGGAACCTCCCTGTGTATGATGATATCCATAGGAACCTTCGACCT 1952  
Db |||||  
Qy 1953 CTGTGAAGATCTCATCTGGAATAAGACCCCTGAGCCACTCAGAGCTCTTACGTTATGC 2012  
Db |||||  
Qy 108 CTGTGAAGATCTCATCTGGAATAAGACCCCTGAGCCACTCAGAGCTCTTACGTTATGC 167  
Db |||||  
Qy 2013 CCAGACTCAAGGCACAGGAGGAGAAAGTCAATTCAGACTGATGAGTGAGAAATGGCCC 2072  
Db |||||  
Qy 168 CCAGACTCAAGGCACAGGAGGAGAAAGTCAATTCAGACTGATGAGTGAGAAATGGCCC 227  
Db |||||  
Qy 2073 TGTCGAAGACGCTTGATGATGCTTGTGAAGGCGATTTGAAACATATTTAGGA 2132  
Db |||||  
Qy 228 TGTCGAAGACGCTTGATGATGCTTGTGAAGGCGATTTGAAACATATTTAGGA 287  
Db |||||  
Qy 2133 TACTGAGAACCCAGGTTAAACCAAAAAATATCCCGACCTCTCAATATAATTTGAAG 2192  
Db |||||  
Qy 288 TACTGAGAACCCAGGTTAAACCAAAAAATATCCCGACCTCTCAATATAATTTGAAG 347  
Db |||||  
Qy 2193 ACCCTGATGAATGGAATGAAATTTGTTGGTATCTTTTGGAGCTGGAATAATGTTTCT 2252  
Db |||||  
Qy 2253 ACCTGAGTTATAAGTCAGCCCGGTTATGAAGAGCTGTTGGCCACCTTATCCCTTT 2312  
Db |||||  
Qy 408 ACCTCAGGTTATAAGTCAGCCCGGTTATGAAGAGCTGTTGGCCACCTTATCCCTTT 467  
Db |||||  
Qy 2313 CATGAAAGAGAGAGAGAAACAGAGTCTTAACGGCAGCTAGAGAGAGAGGCC 2372  
Db |||||  
Qy 468 CATGAAAGAGAGAGAGAAACAGAGTCTTAACGGCAGCTAGAGAGAGAGGCC 527  
Db |||||  
Qy 2373 TTACAGGGCCACCATCGTGTGGCCACTGTTTAAAGGCGCAGCTGCAATAGGCAAGAA 2432  
Db |||||  
Qy 528 TTACAGGGCCACCATCGTGTGGCCACTGTTTAAAGGCGCAGCTGCAATAGGCAAGAA 587  
Db |||||  
Qy 2433 CATAGTGGAGTAGTCTTGGTGGCAATAATTTCCGAGTTATGATTTAGAGTCATGAC 2492  
Db |||||  
Qy 588 CATAGTGGAGTAGTCTTGGTGGCAATAATTTCCGAGTTATGATTTAGAGTCATGAC 647  
Db |||||  
Qy 2493 TCCATGTCAATAGTACTGAAAGCTGCTTGGACCAAGCAGATATAATTTGGCCTGC 2552  
Db |||||  
Qy 648 TCCATGTCAATAGTACTGAAAGCTGCTTGGACCAAGCAGATATAATTTGGCCTGC 707  
Db |||||

RESULT 5

## BX390174

## LOCUS

BX390174 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS00C010Y017 5-PRIME, mRNA sequence.

## ACCESSION

BX390174 BX390174.2 GI:46877980

## VERSION

BX390174.2

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 871)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 8, 2003 this sequence version replaced gi:30463266.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

740.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna/s-CS0BAG0062C05\_CS00515\_1&c=740.f.

Location/Qualifiers

1..871

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS00C010Y017"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 16.8%; Score 651; DB 5; Length 871;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 701; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 671 AGGCAGCCTTTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCGGCTATCT 730

Db |||||

Qy 21 AGGCAGCCTTTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCGGCTATCT 80

Db |||||

Qy 731 TTATTTTCAGGACGATCGTTGATATAAGTGGCGGACTCTTCGGACAGACGAGAGG 790

Db |||||

Qy 81 TTATTTTCAGGACGATCGTTGATATAAGTGGCGGACTCTTCGGACAGACGAGAGG 140

Db |||||

Qy 791 GATTGTTCATCAGCGTCTCATCGAGAACCACTCTGATTGGATTAAATTTGCTTTGG 850

Db |||||

Qy 141 GATTGTTCATCAGCGTCTCATCGAGAACCACTCTGATTGGATTAAATTTGCTTTGG 200

Db |||||

Qy 851 GTGCAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATTTGAAATAATTTGTC 910

Db |||||

Qy 201 GTGCAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATTTGAAATAATTTGTC 260

Db |||||

Qy 911 TCTGTTATCCCAATGCGAGTCTTCCCAACACCTTTGGTGACTATGATGAACCGCTTCTA 970

Db |||||

Qy 261 TCTGTTATCCCAATGCGAGTCTTCCCAACACCTTTGGTGACTATGATGAACCGCTTCTA 320

Db |||||

Qy 971 TGATGGCCCAAGCACCTAAAGATTTCGTATGGATGGCTTCGTCAATATAGTTGGAGAT 1030

Db |||||

Qy 321 TGATGGCCCAAGCACCTAAAGATTTCGTATGGATGGCTTCGTCAATATAGTTGGAGAT 380

Db |||||

Qy 1031 GCTGTGGGTCAACACCGATCATATCAGGGAAATTTGCTGAAGCTGTGAAAAATTTGTAAGC 1090

Db |||||

```

Db 381 GCTGGGTCAACACCAAGATCATATCAGGAAATGCTGAAGCTGTGAAATTTGTAAGC 440
Qy 1091 CTAGAGTCCACCTGCCATCTCTTTGAAGACATATGTTACTGTCTGGTCTAGAGCCCT 1150
Db 441 CTAGAGTCCACCTGCCATCTCTTTGAAGACATATGTTACTGTCTGGTCTAGAGCCCT 500
Qy 1151 TCAGAGTGGACCTGACACCACTTTGTTAACTTGAAGAGGAGGCTGTAAATGTTGCAGAT 1210
Db 501 TCAGAGTGGACCTGACACCACTTTGTTAACTTGAAGAGGAGGCTGTAAATGTTGCAGAT 560
Qy 1211 CAAGCAAGTTGCTAAACTCATATGCGAGCAAACTATGAAGAAGCCTTGTGTGCTGCCA 1270
Db 561 CAAGCAAGTTGCTAAACTCATATGCGAGCAAACTATGAAGAAGCCTTGTGTGCTGCCA 620
Qy 1271 AAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCACATGATGATGCGATGCTAG 1330
Db 621 AAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCACATGATGATGCGATGCTAG 680
Qy 1331 ATGCTCCAGTGCATGACCAAGATTTGCAACTTAATGCTT 1372
Db 681 ATGCTCCAGTGCATGACCAAGATTTGCAACTTAATGCTT 722

RESULT 6
BUI89680
LOCUS 943 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT 7968068 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6170551
5', mRNA sequence.
ACCESSION BUI89680
VERSION BUI89680.1 GI:22703664
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 943)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLML at:
http://image.llnl.gov
Plate: L1AM13537 row: h column: 08
High quality sequence stop: 587.
FEATURES
Location/Qualifiers
source 1..943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6170551"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 16.3%; Score 637; DB 5; Length 943;
Best Local Similarity 99.9%; Pred. No. 5.1e-308;
Matches 687; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2155 CAAAAAATATCCCGACCTCTCAATATATTAATGAAGACCCCTGATGAATGAATGAA 2214
Db 1 CAAAAAATATCCCGACCTCTCAATATATTAATGAAGACCCCTGATGAATGAATGAA 60

```

```

Qy 2215 ATTGTTGGTGTATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAAAGTCAGCC 2274
Db 61 ATTGTTGGTGTATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAAAGTCAGCC 120
Qy 2275 CGGTTTATGAAGAAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAA 2334
Db 121 CGGTTTATGAAGAAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAA 180
Qy 2335 ACCAGAGTGTCTTAACCGCACAGTAGAAGAGAGAGACCTTTACAGGGCACCATCTGCTG 2394
Db 181 ACCAGAGTGTCTTAACCGCACAGTAGAAGAGAGAGACCTTTACAGGGCACCATCTGCTG 240
Qy 2395 GCCACTGTTAAAGGCGAGCTGCGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGC 2454
Db 241 GCCACTGTTAAAGGCGAGCTGCGACGACATAGGCAAGAACATAGTTGGAGTAGTCTTGGC 300
Qy 2455 TGCATAATATTTCCGAGTTATTGATTTAGGAGTCATGACTCCATGTTGATAGATACTGAAA 2514
Db 301 TGCATAATATTTCCGAGTTATTGATTTAGGAGTCATGACTCCATGTTGATAGATACTGAAA 360
Qy 2515 GCTGCTCTTGACCAACAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTG 2574
Db 361 GCTGCTCTTGACCAACAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTG 420
Qy 2575 GATGAATATGATTTTGTGTCAGAGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTG 2634
Db 421 GATGAATATGATTTTGTGTCAGAGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTG 480
Qy 2635 ATTGGAGGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGTCCGAGATACAGT 2694
Db 481 ATTGGAGGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGTCCGAGATACAGT 540
Qy 2695 GCACCTGTAATCCATGTCCTGGACGCTCCAGAGTGTGCTGCTCCAGCTGTTA 2754
Db 541 GCACCTGTAATCCATGTCCTGGACGCTCCAGAGTGTGCTGCTGCTGCTGCTGTTA 600
Qy 2755 GATGAATAATCTAAAGGAGTAAATCTTTGAGGAATCATGGAAGATATGAAGATATTAGA 2814
Db 601 AATGAATAATCTAAAGGAGTAAATCTTTGAGGAATCATGGAAGATATGAAGATATTAGA 660
Qy 2815 CAGGACCATATGAGTCTCTCAAGGAGA 2842
Db 661 CAGGACCATATGAGTCTCTCAAGGAGA 688

RESULT 7
CN414972
LOCUS 698 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000418008177 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN414972
VERSION CN414972.1 GI:47402566
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 698)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murgue, J., Pisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com

```

```
Insert Length: 698 Std Error: 0.00.
FEATURES
  source
    1..698
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
      /clone_lib="GRN ES"
      /note="oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"
ORIGIN
  Query Match 15.9%; Score 623; DB 7; Length 698;
  Best Local Similarity 99.9%; Pred. No. 5.6e-301;
  Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1006 GCCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACGATCATATCAGGAAATT 1065
DB 25 GCCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACGATCATATCAGGAAATT 84
QY 1066 GCTGAAGCTGTGAATAATTTGAAGCTAGAGTTCCATCGTCATGCTTTTGAAGGACAT 1125
DB 85 GCTGAAGCTGTGAATAATTTGAAGCTAGAGTTCCATCGTCATGCTTTTGAAGGACAT 144
QY 1126 ATGTTACTGTCTGGTCTAGAGCCCTTCAGGATTTGACCGTACACCACTTTGTTAAACATT 1185
DB 145 ATGTTACTGTCTGGTCTAGAGCCCTTCAGGATTTGACCGTACACCACTTTGTTAAACATT 204
QY 1186 GGAGAGCCCTGTAAATGTTGAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAAC 1245
DB 205 GGAGAGCCCTGTAAATGTTGAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAAC 264
QY 1246 TATGAAGAAGCCTTGTGTGTTGCCAAAGTCAGAGTGGAAATGGGAGCCAGGTGTTGAT 1305
DB 265 TATGAAGAAGCCTTGTGTGTTGCCAAAGTCAGAGTGGAAATGGGAGCCAGGTGTTGAT 324
QY 1306 GTCACATGATGATGGCATGCTAGATGTTCCAGTGCAGTCAAGTCAAGTCAAGTCAAGTCA 1365
DB 325 GTCACATGATGATGGCATGCTAGATGTTCCAGTGCAGTCAAGTCAAGTCAAGTCAAGTCA 384
QY 1366 ATTGCTTCGAGCCAGACATGCCAAAGTACCTTTGTGTCATCGACTCCTCCAAATTTTGTCT 1425
DB 385 ATTGCTTCGAGCCAGACATGCCAAAGTACCTTTGTGTCATCGACTCCTCCAAATTTTGTCT 444
QY 1426 GTGATTGAAGCTGGGTTAAAGTGTGCCAAGGAAAGTGCATTGTCAATAGCATTTAGTCTG 1485
DB 445 GTGATTGAAGCTGGGTTAAAGTGTGCCAAGGAAAGTGCATTGTCAATAGCATTTAGTCTG 504
QY 1486 AAGGAAGGAGGAGGACGACTTCTTGGAAGAGGCGCAGGAAGATTTAAAGTATGGAGCTGCT 1545
DB 505 AAGGAAGGAGGAGGACGACTTCTTGGAAGAGGCGCAGGAAGATTTAAAGTATGGAGCTGCT 564
QY 1546 ATGGTGGTTCATGGCTTTTATCAAGAGGACGAGCAACAGAAACAGACACAAAAATCAGA 1605
DB 565 ATGGTGGTTCATGGCTTTTATCAAGAGGACGAGCAACAGAAACAGACACAAAAATCAGA 624
QY 1606 GTGTGCACCCGGGCTTACCATTCTGTTGTGAAAAAATCTGGGCTTTTAAATCCAAATGACATT 1665
DB 625 GTGTGCATCCGGGCTTACCATTCTGTTGTGAAAAAATCTGGGCTTTTAAATCCAAATGACATT 684
QY 1666 ATTTTGGACCTTAA 1679
DB 685 ATTTTGGACCTTAA 698
RESULT 8
BP313585 LOCUS BP313585 Sugano cDNA library, EST 17-SEP-2004
DEFINITION BP313585 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
cDNA clone OPR03689, mRNA sequence.
ACCESSION BP313585
```

```
BP313585.1 GI:52242560
EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
  source
    1..583
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="OPR03689"
      /tissue_type="mammary gland"
      /cell_line="OCUB-F"
      /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
      /note="mammary gland tumor"
ORIGIN
  Query Match 13.9%; Score 545; DB 3; Length 583;
  Best Local Similarity 100.0%; Pred. No. 9.3e-262;
  Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2290 GCTGTGGCCACCTTATCCCTTTTCATGAAAGAAAGAGAGAAACACAGAGTCTTAAAC 2349
DB 28 GCTGTGGCCACCTTATCCCTTTTCATGAAAGAAAGAGAGAGAAACACAGAGTCTTAAAC 87
QY 2350 GGACAGTAGAAGAGAGGACCCCTTACAGGGACCATCGTCTGGCCACTGTTTAAAGGC 2409
DB 88 GGACAGTAGAAGAGAGGACCCCTTACAGGGACCATCGTCTGGCCACTGTTTAAAGGC 147
QY 2410 GACGTGCACGACATAGCAGCAACATAGTTGGAGTAGTCTTGGCTGCAATAATTTCCGA 2469
DB 148 GACGTGCACGACATAGCAGCAACATAGTTGGAGTAGTCTTGGCTGCAATAATTTCCGA 207
QY 2470 GTTATTGATTAGGAGTCATGACTCCATGTGATAGATACCTGAAAGCTGCTCTTGACAC 2529
DB 208 GTTATTGATTAGGAGTCATGACTCCATGTGATAGATACCTGAAAGCTGCTCTTGACAC 267
QY 2530 AAAGCAGATATAATTTGGCTGTGAGGACTCATCTCCTCCCTGGATGAATGATTTT 2589
DB 268 AAAGCAGATATAATTTGGCTGTGAGGACTCATCTCCTCCCTGGATGAATGATTTT 327
QY 2590 GTTCCAGGAAATGGAGAGATTAGCTATAGGATTCCTATTGTTGAGGAGGACCAACC 2649
DB 328 GTTCCAGGAAATGGAGAGATTAGCTATAGGATTCCTATTGTTGAGGAGGACCAACC 387
QY 2650 ACTTCAAAAAACCCACACAGCAGATTAAAAATAGCTCCGAGATACAGTGCACCTGTAATCCAT 2709
DB 388 ACTTCAAAAAACCCACACAGCAGATTAAAAATAGCTCCGAGATACAGTGCACCTGTAATCCAT 447
QY 2710 GTCTGGAGCGGCTCCAGAGTGTGGTGTGTTCCTCCAGCTGTTAGATGAATTAAG 2769
DB 448 GTCTGGAGCGGCTCCAGAGTGTGGTGTGTTCCTCCAGCTGTTAGATGAATTAAG 507
QY 2770 GATGAATCTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTATGAG 2829
DB 508 GATGAATCTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTATGAG 567
QY 2830 TCTCT 2834
|||||
```

```
Db 568 TCCTCT 572

RESULT 9
BP336553 Sugano cDNA library, coronary artery smooth muscle cell EST 17-SEP-2004
LOCUS BP336553.1 GI:52266128
DEFINITION Homo sapiens cDNA clone SMR05967, mRNA sequence.
ACCESSION BP336553
VERSION BP336553.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezukui@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR05967"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle cell"
ORIGIN
Query Match 13.5%; Score 529; DB 3; Length 582;
Best Local Similarity 99.8%; Pred. No. 1e-253;
Matches 579; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 680 TGTTCACCTCCAAATCTTTTTCAGGAGAAATATGCTCCCGCCCTATCTTTATTTTCAG 739
Db 3 TGTTCACCTCCAAATCTTTTTCAGGAGAAATATGCTCCCGCCCTATCTTTATTTTCAG 62
Qy 740 GGACGATCGTTGATATAAGTGGCGGACTCTTTCCGACACAGAGAGAGGATTTGTCA 799
Db 63 GGACGATCGTTGATATAAGTGGCGGACTCTTTCCGACACAGAGAGAGGATTTGTCA 122
Qy 800 TCACGGTGTCTCATGGAGAACCTCTGCAATGATTAATTTGCTTTGGGTGCAGCTG 859
Db 123 TCACGGTGTCTCATGGAGAACCTCTGCAATGATTAATTTGCTTTGGGTGCAGCTG 182
Qy 860 AGATGAGACCTTTTATTAATAATTTGAAATGTAACAAGCCCTATGCTCTGTTATC 919
Db 183 AATGAGACCTTTTATTAATAATTTGAAATGTAACAAGCCCTATGCTCTGTTATC 242
Qy 920 CCAATGAGGCTTTCCCAACACCTTTGGTGAATGATGAACAGCCCTTCTATGATGCCA 979
Db 243 CCAATGAGGCTTTCCCAACACCTTTGGTGAATGATGAACAGCCCTTCTATGATGCCA 302
Qy 980 AGCACCTAAGGATTTTCTATGATGCTTGGTCAATATAGTTGGAGGATGCTGTGGGT 1039
Db 303 AGCACCTAAGGATTTTCTATGATGCTTGGTCAATATAGTTGGAGGATGCTGTGGGT 362
Qy 1040 CAACACCATATCATACAGGAAATTTGCTGAAGCTGTGMAAAATTTGTAAGCCTAGAGTTC 1099
Db 363 CAACACCATATCATACAGGAAATTTGCTGAAGCTGTGMAAAATTTGTAAGCCTAGAGTTC 422
Qy 1100 CACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCGTGTAGAGCCCTTCAGGATTG 1159
```

```
Db 423 CACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGTAGAGCCCTTCAGGATTG 482
Qy 1160 GACCGTACACCACTTTTGTAAACATTGGAGAGCGCTGTAAATGTTGCAGGATCAAGGAAGT 1219
Db 483 GACCGTACACCACTTTTGTAAACATTGGAGAGCGCTGTAAATGTTGCAGGATCAAGGAAGT 542
Qy 1220 TTGCTAAACTCATCATGGCAGGAAACTATGACAGACCTT 1259
Db 543 TTGCTAAACTCATCATGGCAGGAAACTATGACAGACCTT 582

RESULT 10
BP337383 Sugano cDNA library, coronary artery smooth muscle cell EST 17-SEP-2004
LOCUS BP337383.1 GI:52266989
DEFINITION Homo sapiens cDNA clone SMR08247, mRNA sequence.
ACCESSION BP337383
VERSION BP337383.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezukui@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR08247"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle cell"
ORIGIN
Query Match 12.8%; Score 501; DB 3; Length 583;
Best Local Similarity 99.8%; Pred. No. 1.2e-239;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 710 AATATGCTCCCGCCCTATCTTTATTTTCAGGAGAGATTTGTCATGAGCGGTCTCATGAGAACCACTCTGCA 769
Db 32 AATATGCTCCCGCCCTATCTTTATTTTCAGGAGAGATTTGTCATGAGCGGTCTCATGAGAACCACTCTGCA 91
Qy 770 TTTCCGACACAGACAGAGAGGATTTGTCATGAGCGGTCTCATGAGAACCACTCTGCA 829
Db 92 TTTCCGACACAGACAGAGAGGATTTGTCATGAGCGGTCTCATGAGAACCACTCTGCA 151
Qy 830 TTGGATTAATTTGCTTTGGGTGCGAGCTGAGATGAGACCTTTATTTGAAATAATTGGAA 889
Db 152 TTGGATTAATTTGCTTTGGGTGCGAGCTGAGATGAGACCTTTATTTGAAATAATTGGAA 211
Qy 890 AATGTACAACAGCCTATGCTCTGTTATCCCAATGAGCTCTTCCCAACACCTTTGGTG 949
Db 212 AATGTACAACAGCCTATGCTCTGTTATCCCAATGAGCTCTTCCCAACACCTTTGGTG 271
Qy 950 ACTATGATGAACCGCTTCTATGATGCGCAACACCTTAAAGGATTTTGTATGATGGCT 1009
Db 272 ACTATGATGAACCGCTTCTATGATGCGCAACACCTTAAAGGATTTTGTATGATGGCT 331
```

```

QY 1010 TGGTCAATATAGTTGGAGGATGCTGTGGTCAACACACGAGTCATATCAGGGAATTTGCTG 1069
Db |||||
332 TGGTCATATATAGTTGGAGGATGCTGTGGTCAACACGAGTCATATCAGGGAATTTGCTG 391
QY 1070 AAGCTGTGAAATTTGTAAGCTAGAGTTCCACCTGCCACCTGCTTTTGAAGGACATATGT 1129
Db |||||
392 AAGCTGTGAAATTTGTAAGCTAGAGTTCCACCTGCCACCTGCTTTTGAAGGACATATGT 451
QY 1130 TACTGTCTGGCTAGAGCCCTTCAGGATTCAGCCTGACACCACTTTGTTAAACATTGGAG 1189
Db |||||
452 TACTGTCTGGCTAGAGCCCTTCAGGATTCAGCCTGACACCACTTTGTTAAACATTGGAG 511
QY 1190 AGCGCTGTAATTTGTCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGACGAACTATG 1249
Db |||||
512 AGCGCTGTAATTTGTCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGACGAACTATG 571
QY 1250 AAGAAGCCTTGT 1261
Db |||||
572 AAGAAGCCTTGT 583

RESULT 11
BP220649 Sugano cDNA library, colon Homo sapiens linear EST 15-SEP-2004
LOCUS BP220649
DEFINITION COL04124, mRNA sequence.
ACCESSION BP220649
VERSION BP220649.1 GI:52093554
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 560)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
FEATURES
source
1..560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL04124"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"

ORIGIN
Query Match 12.7%; Score 497; DB 3; Length 560;
Best Local Similarity 99.8%; Pred. No. 1.3e-237;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2571 CCTGGATGAATGATTTTGTGTCACAGGAATGGAGATTTAGCTATAGGATTTCCATT 2630
Db |||||
13 CCTGGATGAATGATTTTGTGTCACAGGAATGGAGATTTAGCTATAGGATTTCCATT 72
QY 2631 GTTGATTGGAGAGCAACACCTTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATA 2690
Db |||||
73 GTTGATTGGAGAGCAACACCTTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATA 132
QY 2691 CAGTGCACCTGTAATCCATGTCTCTGGAACGCTCAAGAGTGTGGTGTGTGTCTCCAGCT 2750
Db |||||
133 CAGTGCACCTGTAATCCATGTCTCTGGAACGCTCAAGAGTGTGGTGTGTGTCTCCAGCT 192
QY 2751 GTTAGATGAATCTAAAGGATGAATACCTTTTGAGGAATCATGGAAGATATGAAGATAT 2810

```

```

Db |||||
193 GTTAGATGAATCTAAAGGATGAATACCTTTGAGGAATCATGGAAGAAATGAAGATAT 252
QY 2811 TAGACAGGACCATATAGTCTCTCAAGGAGAGAGATATCTTACCCTTAAGTCAAGCCAG 2870
Db |||||
253 TAGACAGGACCATATAGTCTCTCAAGGAGAGAGATATCTTACCCTTAAGTCAAGCCAG 312
QY 2871 AAAAAAGTGGTTTCCAAATGATTTGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTAT 2930
Db |||||
313 AAAAAAGTGGTTTCCAAATGATTTGGCTGTCTGAAACCTCACCCAGTGAAGCCACGTTTAT 372
QY 2931 TGGGACCCAGGCTCTTTGAAGACTATGACCTGCGAGAAGCTGGTGGACTACATTTGACTGGAA 2990
Db |||||
373 TGGGACCCAGGCTCTTTGAAGACTATGACCTGCGAGAAGCTGGTGGACTACATTTGACTGGAA 432
QY 2991 GCCTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGAT 3050
Db |||||
433 GCCTTTCTTTGATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGAT 492
QY 3051 ATTTAACGACAAAACAGTAGTGTGGAGAGCCAGGAAGTCTACGATGATGCCACAATAT 3110
Db |||||
493 ATTTAACGACAAAACAGTAGTGTGGAGAGCCAGGAAGTCTACGATGATGCCACAATAT 552
QY 3111 GCTGAACA 3118
Db |||||
553 GCTGAACA 560

RESULT 12
AL041796 490 bp mRNA linear EST 04-SEP-2003
LOCUS AL041796
DEFINITION DKFP43434B1618_x1_434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL041796
VERSION AL041796.1 GI:5421143
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFP43434B1618) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP43434B1618"
/tissue_type="testis"
/dev stage="adult"
/lab host="DH10B"
/clone_lib="434 (synonym: htes3)"
/notes="vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 12.5%; Score 490; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 4.2e-234;

```





REFERENCE 1 (bases 1 to 582)  
AUTHORS Suzuki Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,  
Mitsushima-Sugano, J., Nakai, K. and Sugano, S.  
TITLE Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
PUBMED 15342556  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="SMR04476"  
/tissue\_type="coronary artery"  
/cell\_type="smooth muscle cell"  
/clone\_lib="Sugano cDNA library, coronary artery smooth  
muscle cell"

ORIGIN  
Query Match 12.2%; Score 480; DB 3; Length 582;  
Best Local Similarity 99.7%; Pred. No. 4.5e-229;  
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 678 CTTGTTTGCACTCCAAATCTTTTGGAGAGAAATATGCTCCCGGCTATCTTTATTTC 737  
Db 1 CTTGTTTGCACTCCAAATCTTTTGGAGAGAAATATGCTCCCGGCTATCTTTATTTC 60  
QY 738 AGGACGATCTGTGTAAGTGGCGGCACTTTCCGACAGACAGAGAGGATTTGT 797  
Db 61 AGGACGATCTGTGTAAGTGGCGGCACTTTCCGACAGACAGAGAGGATTTGT 120  
QY 798 CATCAGCGTCTCATGAGAACCACTCTGCATGGATTAATTTGCTTTGGGTGCAGC 857  
Db 121 CATCAGCGTCTCATGAGAACCACTCTGCATGGATTAATTTGCTTTGGGTGCAGC 180  
QY 858 TGAGATGAGACCTTTTATTTGAATTAATTTGAAAAATGTAACAAGCCTATGCTCTGTTA 917  
Db 181 TGAAATGAGACCTTTTATTTGAATTAATTTGAAAAATGTAACAAGCCTATGCTCTGTTA 240  
QY 918 TCCCAATGCACTCTCCCAACACCTTTGGTGACTATGATGAAGCGCTTCTATGATGC 977  
Db 241 TCCCAATGCACTCTCCCAACACCTTTGGTGACTATGATGAAGCGCTTCTATGATGC 300  
QY 978 CAAGCACCTAAAGGATTTTGTCTATGATGGCTTGTCTCAATATAGTTGGAGGATGCTGTGG 1037  
Db 301 CAAGCACCTAAAGGATTTTGTCTATGATGGCTTGTCTCAATATAGTTGGAGGATGCTGTGG 360  
QY 1038 GTCAACACCCAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAAATTTGTAAGCCCTAGAGT 1097  
Db 361 GTCAACACCCAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAAATTTGTAAGCCCTAGAGT 420  
QY 1098 TCACCTGCACTGCTTTTGAAGGACATATGTTACTGCTGTCTAGAGCCCTTCAGGAT 1157  
Db 421 TCACCTGCACTGCTTTTGAAGGACATATGTTACTGCTGTCTAGAGCCCTTCAGGAT 480  
QY 1158 TGGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAATGTTGAGGATCAAGGAA 1217  
Db 481 TGGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAATGTTGAGGATCAAGGAA 540  
QY 1218 GTTGTCTAACTCATCATGCGAGGAACCTATGAAGAAGCCTT 1259  
Db 541 GTTGTCTAACTCATCATGCGAGGAACCTATGAAGAAGCCTT 582

RESULT 15  
BQ062840  
LOCUS  
DEFINITION AGENCOURT\_6826659 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5924378

5', mRNA sequence.  
BQ062840  
VERSION BQ062840.1 GI:19890017  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Hmo.  
1 (bases 1 to 1043)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: csapbs-remail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM2094 row: 9 column: 03  
High quality sequence start: 242  
High quality sequence stop: 479.

FEATURES  
source  
1..1043  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5924378"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

ORIGIN  
Query Match 11.4%; Score 448; DB 3; Length 1043;  
Best Local Similarity 99.6%; Pred. No. 5.7e-213;  
Matches 668; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2483 GAGTCATGACTCCATGTGATAAGATACCTGAAAGCTGCTTTGACCAAAAGCAGATATA 2542  
Db 118 GAGTCATGACTCCATGTGATAAGATACCTGAAAGCTGCTTTGACCAAAAGCAGATATA 177  
QY 2543 TTGGCCTGTCCAGGACTCATCACTCTCTCCCTGGATGAATGATTTTGTTCGAAGGAAA 2602  
Db 178 TTGGCCTGTCCAGGACTCATCACTCTCTCCCTGGATGAATGATTTTGTTCGAAGGAAA 237  
QY 2603 TGGAGAGATTAGCTATTAAGGATTCATTTGTTGATTTGGAGGACCACTTCAAAAACCC 2662  
Db 238 TGGAGAGATTAGCTATTAAGGATTCATTTGTTGATTTGGAGGACCACTTCAAAAACCC 297  
QY 2663 ACACAGCAGTTTAAATAGCTCCGAGATACAGTGACCTGTGTAATCCATGCTCTGGACGGT 2722  
Db 298 ACACAGCAGTTTAAATAGCTCCGAGATACAGTGACCTGTGTAATCCATGCTCTGGACGGT 357  
QY 2723 CCAAGAGTGTGGTGTGTGTTCCAGCTGTTTAGATGAAAAATCTAAAGATGAATCTTTG 2782  
Db 358 CCAAGAGTGTGGTGTGTGTTCCAGCTGTTTAGATGAAAAATCTAAAGATGAATCTTTG 417  
QY 2783 AGGAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAG-GAG 2841  
Db 418 AGGAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAG-GAG 477

QY 2842 AGGAGATACCTTACCTTAAAGTCAAGCCAGAAAGTGGTTTCCAAATGGATTGGCTGTCT 2901  
 Db 478 AGGAGATACCTTACCTTAAAGTCAAGCCAGAAAGTGGTTTCCAAATGGATTGGCTGTCT 537  
 QY 2902 GAACCTCACCCAGTGAAGCCACGCTTTATTTGGGACCCAGGCTTTTGAAGACTATGACCTG 2961  
 Db 538 GAACTCACCCAGTGAAGCCACGCTTTATTTGGGACCCAGGCTTTTGAAGACTATGACCTG 597  
 QY 2962 CAGAAGCTGGTGGACTACATTTGACTGAGAGCCTTTCTTTGATGTCTGGCAGCTCCGGGGC 3021  
 Db 598 CAGAAGCTGGTGGACTACATTTGACTGAGAGCCTTTCTTTGATGTCTGGCAGCTCCGGGGC 657  
 QY 3022 AAGTACCCGATCGAGGCTCCCAAGATATTTAAACGACAAACAGTAGGTGGAGAGGCC 3081  
 Db 658 AAGTACCCGATCGAGGCTTTCCCAAGATATTTAAACGACAAACAGTAGGTGGAGAGGCC 717  
 QY 3082 AGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTGATTTAGTCAAAAGAACTC 3141  
 Db 718 AGGAAGGTCTACGATGATGCCCAATATGCTGAACACACTGATTTAGTCAAAAGAACTC 777  
 QY 3142 CGGCCCCGGGG 3152  
 Db 778 CGGCCCCGGGG 788

RESULT 16  
 BX485695  
 LOCUS DXFZP686C08248\_r1\_686 (synonym: hlcc3) Homo sapiens cDNA clone  
 DEFINITION DXFZP686C08248\_5', mRNA sequence.  
 ACCSSION BX485695  
 VERSION BX485695.1 GI:31948689  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Homo.

REFERENCE 1 (bases 1 to 626)  
 AUTHORS Ottenwaelder, B., Obermaier, B., Deutscher, B., S., Mewes, H.W.,  
 Weil, B., Amid, C., Oeinger, A., Fobo, G., Han, M. and Wiemann, S.  
 TITLE EST (Ottenwaelder, B., Obermaier, B., Deutscher, B., S., Mewes, H.W.,  
 et al.)

JOURNAL Unpublished (2003)  
 COMMENT Contact: MIPS  
 MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by MediGenomix (Martinsried/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. No 81 sequence  
 available.

This clone (DKFZp686C08248) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
 source  
 1..626  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686C08248"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="686 (synonym: hlcc3)"  
 /note="vector: pTriplex2; Site\_1: sf1A; Site\_2: sf1B;  
 cDNA-collection"

ORIGIN  
 Query Match 11.2%; Score 440; DB 5; Length 626;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-209;  
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1219 TTGTCTAAACTCATCTGCGAGGAACTATGAAGAAAGCCCTTGTTGTGTTGCAAAAGTGCAG 1278  
 Db 1 TTGTCTAAACTCATCTGCGAGGAACTATGAAGAAAGCCCTTGTTGTGTTGCAAAAGTGCAG 60  
 QY 1279 GTGGAAATGGAGCCCGAGGTTTGGATGTCAACATGATGATGGCATGTAGATGGTCCA 1338  
 Db 61 GTGGAAATGGAGCCCGAGGTTTGGATGTCAACATGATGATGGCATGTAGATGGTCCA 120  
 QY 1339 AGTGCATGACCAAGATTTTCAACTTAATTTGCTTCCGAGCCAGACATCGCAAGGTACCT 1398  
 Db 121 AGTGCATGACCAAGATTTTCAACTTAATTTGCTTCCGAGCCAGACATCGCAAGGTACCT 180  
 QY 1399 TTGTGATCGACCTCCCAATTTTGTGCTGATTGAAGCTGGTTAAAGTCTGCTGCAAGGG 1458  
 Db 181 TTGTGATCGACCTCCCAATTTTGTGCTGATTGAAGCTGGTTAAAGTCTGCTGCAAGGG 240  
 QY 1459 AAGTGCATTTCAATAGCATTAGTCTGAAGAAAGGAGGAGCAGCTTTCTTTGGAGAGGCC 1518  
 Db 241 AAGTGCATTTCAATAGCATTAGTCTGAAGAAAGGAGGAGCAGCTTTCTTTGGAGAGGCC 300  
 QY 1519 AGGAAGATTTAAAGTATGAGCTGCTATGCTGATGGTGTATGCTTTTGTATGAAGAGGAGCAG 1578  
 Db 301 AGGAAGATTTAAAGTATGAGCTGCTATGCTGATGGTGTATGCTTTTGTATGAAGAGGAGCAG 360  
 QY 1579 GCAACAGAAACAGACACAAATATCAGAGTGTGCACCGGCGCTACCATCTGCTTGTGAAA 1638  
 Db 361 GCAACAGAAACAGACACAAATATCAGAGTGTGCACCGGCGCTACCATCTGCTTGTGAAA 420  
 QY 1639 AAAGTGGGCTTTAATCCAAA 1658  
 Db 421 AAAGTGGGCTTTAATCCAAA 440

RESULT 17  
 CV334327  
 LOCUS CV334327 404 bp mRNA linear EST 24-SEP-2004  
 DEFINITION IL3-WT0280-250101-421-E07 NT0280 Homo sapiens cDNA, mRNA sequence.  
 ACCSSION CV334327  
 VERSION CV334327.1 GI:52657541  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 404)  
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 PUBMED 10737800

COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. http://www.ludwig.org.br.  
 Location/Qualifiers  
 1..404  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="NT0280"

FEATURES  
 source

/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 10.3%; Score 404; DB 7; Length 404;  
Best Local Similarity 100.0%; Pred. No. 7.4e-191;  
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2508 ACTGAAGCTCTCTTGACCAAGCAGATATATATGGCCCTGTCAGGACTCATCACTCC 2567  
DB 1 ACTGAAGCTCTCTTGACCAAGCAGATATATATGGCCCTGTCAGGACTCATCACTCC 60

QY 2568 TTCCCTGATGAATGATTTTGTTCACCAAGAAATGAGAGATTAGTATAAGGATTC 2627  
DB 61 TTCCCTGATGAATGATTTTGTTCACCAAGAAATGAGAGATTAGTATAAGGATTC 120

QY 2628 ATTGTTGATTGGAGGAGCAACACATTTCAAAAACCCACACAGAGTTAAATAGCTCCGAG 2687  
DB 121 ATTGTTGATTGGAGGAGCAACACATTTCAAAAACCCACACAGAGTTAAATAGCTCCGAG 180

QY 2688 ATACAGTGCACTGTAAATCCATGTCCTGGACGCGTCCAGAGTGTGGTGTGTCCCA 2747  
DB 181 ATACAGTGCACTGTAAATCCATGTCCTGGACGCGTCCAGAGTGTGGTGTGTCCCA 240

QY 2748 GCTGTTAGATGAATATCTAAAGGATGAATCTTTGAGGAATCATGGAAGAAATATGAAGA 2807  
DB 241 GCTGTTAGATGAATATCTAAAGGATGAATCTTTGAGGAATCATGGAAGAAATATGAAGA 300

QY 2808 TATTAGACAGGACCACTTATGAGTCTCTCAAGAGAGGAGATCTTACCCTTAAGTCAAGC 2867  
DB 301 TATTAGACAGGACCACTTATGAGTCTCTCAAGAGAGGAGATCTTACCCTTAAGTCAAGC 360

QY 2868 CAGAAAAGTGGTTTCCAAATGGATGCTGCTCAACCTCACC 2911  
DB 361 CAGAAAAGTGGTTTCCAAATGGATGCTGCTCAACCTCACC 404

RESULT 18  
LOCUS BQ960377  
DEFINITION AGENCOURT\_8946185 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6462473 5', mRNA sequence.  
ACCESSION BQ960377  
VERSION BQ960377.1 GI:22375855  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 937)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLCM2652 row: g column: 18  
High quality sequence stop: 568.  
Location/Qualifiers  
1..937  
/organism="Homo sapiens"

FEATURES  
source

/mol type="mRNA"  
/db xref="taxon:9606"  
/clone="IMAGE:6462473"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_101"  
/notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 9.3%; Score 363; DB 5; Length 937;  
Best Local Similarity 99.4%; Pred. No. 3.2e-170;  
Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3361 TTTGCCGTTCCCTGCTTTGGGTTAGAGAGCTGAGCAAGCCCTATGAGGATGATGGTAC 3420  
DB 1 TTTGCCGTTCCCTGCTTTGGGTTAGAGAGCTGAGCAAGCCCTATGAGGATGATGGTAC 60

QY 3421 GACTACAGCAGCATCATGTTCAAGCGCTCGGGGACCGGCTGGCAGAGGCTTTGCAGAA 3480  
DB 61 GACTACAGCAGCATCATGTTCAAGCGCTCGGGGACCGGCTGGCAGAGGCTTTGCAGAA 120

QY 3481 GAGTCCATGAAGAGTTCCCGGAGAACTGTGGGCCCTACTGTGGCAGTGAGCAGCTGGAC 3540  
DB 121 GAGTCCATGAAGAGTTCCCGGAGAACTGTGGGCCCTACTGTGGCAGTGAGCAGCTGGAC 180

QY 3541 GTCCGACACTGCGGAAGTTGCGGTACAAGGGCATCCGCCGCTCTGCTGCTACCCGAGC 3600  
DB 181 GTCCGACACTGCGGAAGTTGCGGTACAAGGGCATCCGCCGCTCTGCTGCTACCCGAGC 240

QY 3601 CAGCCGACACACACCCGAGAGCTCACCATGTGGAGACTCCGACACATCGAGCAGCTCTACA 3660  
DB 241 CAGCCGACACACACCCGAGAGCTCACCATGTGGAGACTCCGACACATCGAGCAGCTCTACA 300

QY 3661 GGCAATTAGGTTTAAACAGAAATCATTTAGCAATGGCACTGCTCTTCCAGCAGTCTCAGGCTCTAC 3720  
DB 301 GGCAATTAGGTTTAAACAGAAATCATTTAGCAATGGCACTGCTCTTCCAGCAGTCTCAGGCTCTAC 360

QY 3721 TTCTCCCAATTTGAAGTCCAAATATTTTCTGTGGGGAAGATTTCCAGAGGATCAGGTTGAG 3780  
DB 361 TTCTCCCAATTTGAAGTCCAAATATTTTCTGTGGGGAAGATTTCCAGAGGATCAGGTTGAG 420

QY 3781 GATTATGATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTGGACCCCAT 3840  
DB 421 GATTATGATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGCTTGGACCCCAT 480

QY 3841 TTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3876  
DB 481 TTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 516

RESULT 19  
LOCUS BE879762  
DEFINITION BE879762 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3893729 5', mRNA sequence.  
ACCESSION BE879762  
VERSION BE879762.1 GI:10328538  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 874)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

**JOURNAL COMMENT**  
 Unpublished (1999)  
 Contact: Robert Straubers, Ph.D.  
 Email: c9pbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTF/Gazdar  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9682 row: d column: 18  
 High quality sequence stop: 561.

**FEATURES**  
 source  
 1..874  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3893729"  
 /tissue\_type="large cell carcinoma, undifferentiated"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_69"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."

**ORIGIN**  
 Query Match 8.9%; Score 350; DB 2; Length 874;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-163;  
 Matches 450; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3377 TTGGGGTAGAAGAGCTGAGCAGAGCCCTATGAGGATGATGAGCAGCTACAGCAGCATCA 3436  
 Db 77 TTGGGGTAGAAGAGCTGAGCAGAGCCCTATGAGGATGATGAGCAGCTACAGCAGCATCA 136  
 Qy 3437 TGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTGAGAGAGCTCCATGAAAGAG 3496  
 Db 137 TGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTGAGAGAGCTCCATGAAAGAG 196  
 Qy 3497 TTGCCCGAGAACTGTGGGCTTACTGTGGCAGTACGAGCTGGAGCTCGCAGACCTCGGAA 3556  
 Db 197 TTGCCCGAGAACTGTGGGCTTACTGTGGCAGTACGAGCTGGAGCTCGCAGACCTCGGAA 256  
 Qy 3557 GGTGGCGGTACAGGGGATCGCCCGCTCTGGCTACCCAGCAGCCGACACACCG 3616  
 Db 257 GGTGGCGGTACAGGGGATCGCCCGCTCTGGCTACCCAGCAGCCGACACACCG 316  
 Qy 3617 AGAAGCTCACCATGTGAGACTTCGAGACATCGAGCAGTCTACAGGCAATTAGGTTACAG 3676  
 Db 317 AGAAGCTCACCATGTGAGACTTCGAGACATCGAGCAGTCTACAGGCAATTAGGTTACAG 376  
 Qy 3677 AATCATTAGCAATGGCACTTCTTACAGCAGTCTCAGGCTTACTTCTCCAATTGAAAT 3736  
 Db 377 AATCATTAGCAATGGCACTTCTTACAGCAGTCTCAGGCTTACTTCTCCAATTGAAAT 436  
 Qy 3737 CCAAAATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTTGAGGATTTGCAATTGAGGA 3796  
 Db 437 CCAAAATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTTGAGGATTTGCAATTGAGGA 496  
 Qy 3797 AGACATATCTGTGGCTGAGGTTGAGAAATGG 3828  
 Db 497 AGACATATCTGTGGCTGAGGTTGAGAAATGG 528

**RESULT 20**  
 BP268350  
 LOCUS BP268350 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone  
 DEFINITION JTH09813, mRNA sequence.  
 ACCESSION BP268350  
 VERSION BP268350.1 GI:52183582  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 579)  
 AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.  
 TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions  
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
 PUBMED 15342556  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: yusuzuki@ims.u-tokyo.ac.jp.

**FEATURES**  
 source  
 1..579  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="JTH09813"  
 /tissue\_type="thyroid"  
 /cell\_line="JTH"  
 /clone\_lib="Sugano cDNA library, thyroid JTH"  
 /note="thyroid tumor"

**ORIGIN**  
 Query Match 8.5%; Score 332; DB 3; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-154;  
 Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 525 TGTGTCCCCATCTGTGAAAGCGCGGATATAGGAACATCACATTTGATGAGCTTGTGA 584  
 Db 16 TGTGTCCCCATCTGTGAAAGCGCGGATATAGGAACATCACATTTGATGAGCTTGTGA 75  
 Qy 585 AGCATACCAAGACAGAGCCCAAGGACTCTGGATGGCGGGTGTATCTTACTCATGA 644  
 Db 76 AGCATACCAAGACAGAGCCCAAGGACTCTGGATGGCGGGTGTATCTTACTCATGA 135  
 Qy 645 AACTATTTTGTACTGCCAATGCCAAGGAGCTTGTGTGCACTCCAAATCTTTTGA 704  
 Db 136 AACTATTTTGTACTGCCAATGCCAAGGAGCTTGTGTGCACTCCAAATCTTTTGA 195  
 Qy 705 GGAGAAATATGCTCCCGGCTATCTTTATTTTCAGGACGATCGTTGATAAAGTGGCG 764  
 Db 196 GGAGAAATATGCTCCCGGCTATCTTTATTTTCAGGACGATCGTTGATAAAGTGGCG 255  
 Qy 765 GACTCTTCCGACAGACAGGAGGATTTGTTCATCAGGCTGTCTCATGAGAACCACT 824  
 Db 256 GACTCTTCCGACAGACAGGAGGATTTGTTCATCAGGCTGTCTCATGAGAACCACT 315  
 Qy 825 CTGCTTGGATTAATTTGCTTTGGTGCAG 856  
 Db 316 CTGCTTGGATTAATTTGCTTTGGTGCAG 347

**RESULT 21**  
 BX470290  
 LOCUS BX470290 527 bp mRNA linear EST 04-SEP-2003  
 DEFINITION DKFZP779J129\_r1 779 (synonym: lnccl) Homo sapiens cDNA clone  
 DKFZP779J129 5', mRNA sequence.  
 ACCESSION BX470290  
 VERSION BX470290.1 GI:32023018  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 527)  
 AUTHORS Bloeker, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.

TITLE EST (Bloeker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS  
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No s1 sequence available.  
This clone (DKFZp779J129) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1. 527  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp779J129"  
/tissue\_type="liver"  
/dev\_stage="fetal"  
/lab\_host="DH10B"  
/clone\_lib="779 (synonym: hnccl)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiA; Site\_2: SfiIB"

FEATURES  
source  
Query Match 8.2%; Score 323; DB 5; Length 527;  
Best Local Similarity 99.2%; Pred. No. 4e-150;  
Matches 523; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2878 GGTTCCTCAATGGATGGCTGTGAACCTCACCAGTGAAGCCACGTTATTGGGACC 2937  
Db 1 GGTTCCTCAATGGATGGCTGTGAACCTCACCAGTGAAGCCACGTTATTGGGACC 60  
QY 2938 CAGGCTTTTGAAGACTATGACTCGCAGAACTGGTGGACTACATTGACTGGAAGCTTTTC 2997  
Db 61 CAGGCTTTTGAAGACTATGACTCGCAGAACTGGTGGACTACATTGACTGGAAGCTTTTC 120  
QY 2998 TTTGATGTCTGGCAGCTCCGGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTAAC 3057  
Db 121 TTTGATGTCTGGCAGCTCCGGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTAAC 180  
QY 3058 GACAAAACAGTAGTGGAGAGCCAGGAAGGTCTACGATGATGCCACAAATATGCTGAAC 3117  
Db 181 GACAAAACAGTAGTGGAGAGCCAGGAAGGTCTACGATGATGCCACAAATATGCTGAAC 240  
QY 3118 ACACGTATTAGTCAAAAGAAACTCCGGCCCGGGGTGTGGTTCTGGCCAGCAGCAG 3177  
Db 241 ACACGTATTAGTCAAAAGAAACTCCGGCCCGGGGTGTGGTTCTGGCCAGCAGCAG 300  
QY 3178 AGTATCCAAGAGCAGATTCACCTGTACGAGAGGCTGTGTGCCCGCAGGCTGCAGAGCCC 3237  
Db 301 AGTATCCAAGAGCAGATTCACCTGTACGAGAGGCTGTGTGCCCGCAGGCTGCAGAGCCC 360  
QY 3238 ATAGCCACTTTCATGGGTAAAGCAACAGGCTGAGAAAGGACTGTGCCAGCAGCAGGACCA 3297  
Db 361 ATAGCCACTTTCATGGGTAAAGCAACAGGCTGAGAAAGGACTGTGCCAGCAGCAGGACCA 420  
QY 3298 TACTACTGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGAGACTTACTGGGC 3357  
Db 421 TACTACTGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGAGACTTACTGGGC 480  
QY 3358 CTGTTTGGCGTTGCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGGCCCTA 3404  
Db 481 CTGTTTGGCGTTGCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGGCCCTA 527

RESULT 22  
AW277117/c 523 bp mRNA linear EST 03-JAN-2000  
LOCUS xp63a10.x1 NCI\_CGAP\_Ov39 Homo sapiens cDNA clone IMAGE:2745018 3'  
DEFINITION

similar to SW:METH HUMAN Q99707  
5-METHYLTETRAHYDROFOLATE--HOMOCYSTEINE METHYLTRANSFERASE ;, mRNA  
sequence.  
AW277117 GI:6664147  
EST.  
KEYWORDS Homo sapiens (human)  
SOURCE ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 523)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-x@mail.nih.gov  
Tissue Procurement: Andrew Berchuck M.D., John Gillespie M.D.,  
Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
cDNA Library Arrayed by: Greg Lemmon, Ph.D.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found  
Seq primer: -400P from Gibco  
High quality sequence stop: 421.

FEATURES  
source  
1. 523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2745018"  
/sex="female"  
/tissue\_type="papillary serous ovarian metastasis"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Ov39"  
/note="Organ: ovary; Vector: pAMP10; cDNA made by oligo-dT  
priming. Non-directionally cloned into the UDG sites of  
pAMP10. Size-selected on agarose gel, average insert  
size 500 bp. Primary library; non-amplified. cDNA  
Library Preparation: David B. Krizman, Ph.D (NCI).  
Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."

ORIGIN  
Query Match 8.2%; Score 321; DB 1; Length 523;  
Best Local Similarity 99.5%; Pred. No. 4.1e-149;  
Matches 421; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1328 TAGATGTGCCAAGTGCATGACCAAGATTTTGCAACTTAATGCTTCGAGCCAGACATCG 1387  
Db 432 TAGATGTGCCAAGTGCATGACCAAGATTTTGCAACTTAATGCTTCGAGCCAGACATCG 373  
QY 1388 CAAAGGTACCTTTGTGCATCGACTCCTCCAATTTTGTGTAATGAAGCTGGGTAAAGT 1447  
Db 372 CAAAGGTACCTTTGTGCATCGACTCCTCCAATTTTGTGTAATGAAGCTGGGTAAAGT 313  
QY 1448 GCTGCCAAGGGAAGTGCATTTGTCAATAGCATTAGTCTGAGGAGGAGGAGGACGACTTCT 1507  
Db 312 GCTGCCAAGGGAAGTGCATTTGTCAATAGCATTAGTCTGAGGAGGAGGAGGACGACTTCT 253  
QY 1508 TGGAGAAGGCCAGCAAGATTAAGAAAGTATGAGCTGCTATGCTGCTCATGCTTTTGATG 1567  
Db 252 TGGAGAAGGCCAGCAAGATTAAGAAAGTATGAGCTGCTATGCTGCTCATGCTTTTGATG 193  
QY 1568 AAGAAGGACAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCACCCGGGCCCTACCATC 1627  
Db 192 AAGAAGGACAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCACCCGGGCCCTACCATC 133



```

Qy 1628 TGCTGTGAAAAAATGGGCTTTAATCCAAATGACATATTTTGGCCCTAATATCTTAA 1687
Db 132 TGCTGTGAAAAAATGGGCTTTAATCCAAATGACATATTTTGGCCCTAATATCTTAA 73
Qy 1688 CCATTGGGACTGGAAATGGAGGAACACAACTTGATGCCATTAAATTTTATCCATCAACAA 1747
Db 72 CCATTGGGACTGGAAATGGAGGAACACAACTTGATGCCATTAAATTTTATCCATCAACAA 13
Qy 1748 AAG 1750
Db 12 AAG 10

RESULT 23
BQ364647
LOCUS MR2-SN0005-150500-002-d12 SN0005 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ364647
VERSION BQ364647.1 GI:21037237
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 506)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-SN0005-
150500-002-d12&t3=2000-05-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 505.
FEATURES
Location/Qualifiers
source 1..506
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0005"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 8.0%; Score 315; DB 3; Length 506;
Best Local Similarity 99.5%; Pred. No. 4.2e-146;
Matches 415; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 877 GAAATAATGGAAAAATGTACAAAGCCCTATCTCTGTTATCCCAATGCAAGGTCCTCCC 936
|||||

```

```

Db 55 GAAATAATGGAAAAATGTACAAAGCCCTATCTCTGTTATCCCAATGCAAGGTCCTCCC 114
Qy 937 AACACCTTTGGTCACTATGATGAAAGCCCTTCTATGATGCCCAAGCACCTTAAAGATTTT 996
|||||
Db 115 AACACCTTGGTCACTATGATGAAAGCCCTTCTATGATGCCCAAGCACCTTAAAGATTTT 174
|||||
Qy 997 GCTATGGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACCAAGATCATATC 1056
|||||
Db 175 GCTATGGATGGCTTGGTCAATATAGTTGGAGGATGCTGTGGGTCAACACCAAGATCATATC 234
|||||
Qy 1057 AGGGAATTTGCTGAAGCTGTGAAAAATTTGAAGCCTAGAGTTCCACCTGCCACCTGCTTTT 1116
|||||
Db 235 AGGGAATTTGCTGAAGCTGTGAAAAATTTGAAGCCTAGAGTTCCACCTGCCACCTGCTTTT 294
|||||
Qy 1117 GAAGGACATATGTTACTGCTGTAGAGCCCTTCAAGGATGGACCGTACACCAACTTA 1176
|||||
Db 295 GAAGGACATATGTTACTGCTGTAGAGCCCTTCAAGGATGGACCGTACACCAACTTA 354
|||||
Qy 1177 GTTAACATTGGAGAGCGCTGTAATGTTGCAGGATCAAGGAAGTTTGCTAACTCATATG 1236
|||||
Db 355 GTTAACATTGGAGAGCGCTGTAATGTTGCAGGATCAAGGAAGTTTGCTAACTCATATG 414
|||||
Qy 1237 GCAGGAACTATGAAGAGCCTTGTGTGTTGCCAAAGTCAGGTGGAAATGGGAGCC 1293
|||||
Db 415 GCAGGAACTATGAAGAGCCTTGTGTGTTGCCAAAGTCAGGTGGAAATGGGAGCC 471

RESULT 24
BQ258321
LOCUS BQ258321
DEFINITION 602379766F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4510409 5',
mRNA sequence.
ACCESSION BQ258321
VERSION BQ258321.1 GI:12768137
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 997)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-kemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM10392 row: c column: 18
High quality sequence stop: 640.
FEATURES
Location/Qualifiers
source 1..997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4510409"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 8.0%; Score 315; DB 2; Length 997;
Best Local Similarity 100.0%; Pred. No. 4.4e-146;

```

```
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2229 TTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATATAAGTCAGCCGGGTTATGAAGAA 2288
Db 78 TTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATATAAGTCAGCCGGGTTATGAAGAA 137
QY 2289 GGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAACAGAGAGTCTTAA 2348
Db 138 GGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAGAACAGAGAGTCTTAA 197
QY 2349 CGGCACAGTGAAGAAGAGACCTTACCGAGGACCAATCGTGTGGCCACTGTTAAAGG 2408
Db 198 CGGCACAGTGAAGAAGAGAGCCCTTACCGAGGACCAATCGTGTGGCCACTGTTAAAGG 257
QY 2409 CGAGTGCACACATAGGCAGAGACATAGTTGGAGTAGTCCCTGGCTGCATATTTCCG 2468
Db 258 CGAGTGCACACATAGGCAGAGACATAGTTGGAGTAGTCCCTGGCTGCATATTTCCG 317
QY 2469 AGTTATTGATTAGGAGTCATGACTCCATGTGATAAGATATGAAAGCTGCTTTGACCA 2528
Db 318 AGTTATTGATTAGGAGTCATGACTCCATGTGATAAGATATGAAAGCTGCTTTGACCA 377
QY 2529 CAAAGCAGATATAAT 2543
Db 378 CAAAGCAGATATAAT 392
```

```
RESULT 25
LOCUS BX332266 961 bp mRNA linear EST 08-APR-2004
DEFINITION BX332266 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC010Y017 5-PRIME, mRNA sequence.
ACCESSION BX332266
VERSION BX332266.2 GI:46280008
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 961)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30343109.
Contact: Genoscope
Genoscope - Centre National de Sequençage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
740.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC010AH09QPI&c=740.f.
```

```
FEATURES
source
1..961
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC010Y017"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

## ORIGIN

Query Match 8.0%; Score 314; DB 5; Length 961;  
Best Local Similarity 100.0%; Pred. No. 1.4e-145;

```
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 389 GCCTTGAACACTTGGCCCTACCGATGAACATGTCTCTGCAGAGTGCCAGAAAAAGCTG 448
Db 526 GCCTTGAACACTTGGCCCTACCGATGAACATGTCTCTGCAGAGTGCCAGAAAAAGCTG 585
QY 449 CCAGAGGAGGTAACTCTCTCCAGACAGGAATTAAAGAGTTTGTGGCAGGGGCTCTGGGTCCGA 508
Db 586 CCAGAGGAGGTAACTCTCTCCAGACAGGAATTAAAGAGTTTGTGGCAGGGGCTCTGGGTCCGA 645
QY 509 CTAATAAGACACTCTCTGTGTGCCCATCTGTGGAAGCCGCGATATATAGGAACATCAT 568
Db 646 CTAATAAGACACTCTCTGTGTGCCCATCTGTGGAAGCCGCGATATATAGGAACATCAT 705
QY 569 TTGATGAGCTTGTGGAAGCATACCAAGCAGAGGCCAAAGGACTTCTCGATGGCGGGTTG 628
Db 706 TTGATGAGCTTGTGGAAGCATACCAAGCAGAGGCCAAAGGACTTCTCGATGGCGGGTTG 765
QY 629 ATATCTTACTCATTTGAAACTATTTTTCATATCTGCCAATGCCAAGCGACCTTGTTTGCAC 688
Db 766 ATATCTTACTCATTTGAAACTATTTTTCATATCTGCCAATGCCAAGCGACCTTGTTTGCAC 825
QY 689 TCCAAATCTTTTT 702
Db 826 TCCAAATCTTTTT 839
```

```
RESULT 26
LOCUS CV344680 574 bp mRNA linear EST 24-SEP-2004
DEFINITION MR1-HN0069-020101-013-h11 HN0069 Homo sapiens cDNA, mRNA sequence.
ACCESSION CV344680
VERSION CV344680.1 GI:52667894
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 574)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. http://www.ludwig.org.br.
Location/Qualifiers
1..574
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HN0069"
/notes="Organ: head normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR [U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research]
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
```

Project. http://www.ludwig.org.br.

## FEATURES

source

1..574  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HN0069"

Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR [U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research]  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under

```

ORIGIN
Query Match          7.6%; Score 298; DB 7; Length 574;
Best Local Similarity 99.7%; Pred. No. 1.5e-137;
Matches 348; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2367 GGACCTTACCAGGCGCCATCGTGGCGGACCTGTTAAAGGCGACGTGACGACATAGG 2426
DB 393 GGACCTTACCAGGCGCCATCGTGGCGGACCTGTTAAAGGCGACGTGACGACATAGG 334
QY 2427 CAAGAACATAGTTCGAGTAGTTCCTGGCTGCAATAATTTCCGAGTATTGATTTAGAGT 2486
DB 333 CAAGAACATAGTTCGAGTAGTTCCTGGCTGCAATAATTTCCGAGTATTGATTTAGAGT 274
QY 2487 CATGACTCCATGTGATAGATGACTGAAAGCTGCTCTTGACCAACAAGCAGATATAATTGG 2546
DB 273 CATGACTCCATGTGATAGATGACTGAAAGCTGCTCTTGACCAACAAGCAGATATAATTGG 214
QY 2547 CCTGTGAGGACTCATCACTCCCTCCCTGGATGAAATGATTTTGTGCAAGGAAATGGA 2606
DB 213 CCTGTGAGGACTCATCACTCCCTCCCTGGATGAAATGATTTTGTGCAAGGAAATGGA 154
QY 2607 GAGATTAGCTATAGGATTCATTTGTTGATGGAGGAGCACCACTTCAAAAACCCACAC 2666
DB 153 GAGATTAGCTATAGGATTCATTTGTTGATGGAGGAGCACCACTTCAAAAACCCACAC 94
QY 2667 AGCAGTTAAATAGCTCCGAGATACAGTGCACCTGTAATCCATGCTCCTG 2715
DB 93 AGCAGTTAAATAGCTCCGAGATACAGTGCACCTGTAATCCATGCTCCTG 45

RESULT 27
BP336596
LOCUS BP336596 Sugano cDNA library, coronary artery smooth muscle cell
DEFINITION Homo sapiens cDNA clone SMR06067, mRNA sequence.
ACCESSION BP336596
VERSION BP336596
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J.,
Mizushima-Sugano J., Nakai K. and Sugano S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
FEATURES
source
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR06067"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
muscle cell"

ORIGIN
Query Match          7.3%; Score 287; DB 3; Length 581;
Best Local Similarity 99.5%; Pred. No. 5.1e-132;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 697 CTTTGTGAGGAGAAATATGCTCCCGGCCCTATCTTTATTTTCAGGAGCATGTTGATAAA 756
DB 19 CTTTGTGAGGAGAAATATGCTCCCGGCCCTATCTTTATTTTCAGGAGCATGTTGATAAA 78
QY 757 AGTGGGCGGACTCTTTCCGACAGACAGAGAGAGGATTTGTCATCAGCGTGTCTCATGGA 816
DB 79 AGTGGGCGGACTCTTTCCGACAGACAGAGAGAGGATTTGTCATCAGCGTGTCTCATGGA 138
QY 817 GAACCACTCTGCAATGGAATTAATTTGCTTTGGGTGCGAGTCAGATGAGACCTTTTATT 876
DB 139 GAACCACTCTGCAATGGAATTAATTTGCTTTGGGTGCGAGTCAGATGAGACCTTTTATT 198
QY 877 GAATTAATTTGGAATAATGTAACAAGCCTATGCTCTCTGTTATCCCAATGCAAGTCTTCCC 936
DB 199 GAATTAATTTGGAATAATGTAACAAGCCTATGCTCTCTGTTATCCCAATGCAAGTCTTCCC 258
QY 937 AACACCTTTGGTGACTATGATGAAGCGCTTCTTATGATGCCCAAGCACTTAAAGGATTTT 996
DB 259 AACACCTTTGGTGACTATGATGAAGCGCTTCTTATGATGCCCAAGCACTTAAAGGATTTT 318
QY 997 GCTATGATGGCTTGGTCAATATAGTTGAGAGATGCTGTGGGTCAACACAGATCATATC 1056
DB 319 GCTATGATGGCTTGGTCAATATAGTTGAGAGATGCTGTGGGTCAACACAGATCATATC 378
QY 1057 AGGGAATTCCTGAGCTGTGAAAATTG 1085
DB 379 AGGGAATTCCTGAGCTGTGAAAATTG 407

RESULT 28
BX498595
LOCUS DXFZp779G0140 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION DXFZp779G0140 5', mRNA sequence.
ACCESSION BX498595
VERSION BX498595.1 GI:32016249
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 587)
AUTHORS Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
Fobo G., Han M. and Wiemann S.
TITLE EST (Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GSF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No 81 sequence available.
This clone (DKFZp779G0140) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
FEATURES
source
1..587
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp779G0140"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIB"

ORIGIN

```

```
Query Match      6.8%; Score 267; DB 5; Length 587;
Best Local Similarity 99.7%; Pred. No. 5.9e-122;
Matches 317; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3560 TGCAGTCAAGGGGATCCGCCCGCTCTGCTACCCAGCAGCCGACACACCCGAGA 3619
Db 140 TGCAGTCAAGGGGATCCGCCCGCTCTGCTACCCAGCAGCCGACACACCCGAGA 199

Qy 3620 AGCTCACCATGTGAGACTCCGACACATCGAGCAGTCTACAGGCAATAGGTTACAGAA 3679
Db 200 AGCTCACCATGTGAGACTCTGACACATCGAGCAGTCTACAGGCAATAGGTTACAGAA 259

Qy 3680 CATTAGCAATGGCAGCTCTCAGCAGTCTCAGGCTCTACTCTCCAAATTGAAGTCCA 3739
Db 260 CATTAGCAATGGCAGCTCTCAGCAGTCTCAGGCTCTACTCTCCAAATTGAAGTCCA 319

Qy 3740 AATATTTTGTGTGGGGAAGATTTCCAGGATCAGGTTGAGGATATGCAATTGAGGAAGA 3799
Db 320 AATATTTTGTGTGGGGAAGATTTCCAGGATCAGGTTGAGGATATGCAATTGAGGAAGA 379

Qy 3800 ACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCCATTTTGGGATATGATACAGACT 3859
Db 380 ACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCCATTTTGGGATATGATACAGACT 439

Qy 3860 AACTTTTTTTTTTTTTTTT 3877
Db 440 AACTTTTTTTTTTTTTTTT 457

RESULT 29
CR629973
LOCUS      CR629973      623 bp      mRNA      linear      EST 11-AUG-2004
DEFINITION DKFZp46910922_r1_469 (synonym: pkid1) Pongo pygmaeus CDNA clone
ACCESSION  CR629973
VERSION    CR629973.1 GI:51126053
KEYWORDS  EST.
SOURCE    Pongo pygmaeus (orangutan)
ORGANISM  Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pongo.
REFERENCE  1 (bases 1 to 623)
AUTHORS   Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Pobo, G., Han, M. and Wiemann, S.
TITLE     Pongo pygmaeus mRNA (Bahr, A., Lauber, J., Mewes, H.W., et al.)
JOURNAL   Unpublished (2004)
COMMENT   Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen
(Hilden/Germany) within the CDNA sequencing consortium of the
German Genome Project. This clone (DKFZp46910922) is available at
the RZPD Deutsche Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp46910922
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES             source
    Location/Qualifiers
        1..623
            /organism="Pongo pygmaeus"
            /mol_type="mRNA"
            /db_xref="taxon:9600"
            /clone="DKFZp46910922"
            /tissue_type="kidney"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="469 (synonym: pkid1)"
            /note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"

ORIGIN
Query Match      6.4%; Score 251; DB 3; Length 593;
Best Local Similarity 100.0%; Pred. No. 6.5e-114;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Query Match      6.8%; Score 267; DB 7; Length 623;
Best Local Similarity 99.3%; Pred. No. 5.9e-122;
Matches 417; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2622 GATTCCATTGTTGATTTGGAGGACCAACCACTTCAAAAACCCACACAGCAGTTAAATAGC 2681
Db 71 GATTCCATTGTTGATTTGGAGGACCAACCACTTCAAAAACCCACACAGCAGTTAAATAGC 130

Qy 2682 TCCGAGATACAGTGACCTCTTAATCCATGTCCTGGAGCGCTCCAAAGAGTGTGGTGGTG 2741
Db 131 TCCGAGATACAGTGACCTCTTAATCCATGTCCTGGAGCGCTCCAAAGAGTGTGGTGGTG 190

Qy 2742 TTCCAGCTCTTAGATGAAATCTAAAGATGAATACTTTGAGGAAATCATGGAAGAATA 2801
Db 191 TTCTCAGCTCTTAGATGAAATCTAAAGATGAATACTTTGAGGAAATCATGGAAGAATA 250

Qy 2802 TGAAGATATTAGACAGACCAATTATGAGTCTCTCAAGGAGAGGAGATACCTTACCCTTAAG 2861
Db 251 TGAAGATATTAGACAGACCAATTATGAGTCTCTCAAGGAGAGGAGATACCTTACCCTTAAG 310

Qy 2862 TCAAGCCAGAAAAGTGGTTTCCAAATGGATGGCTCTCTGAAACCTCACCCAGTGAAGCC 2921
Db 311 TCAAGCCAGAAAAGTGGTTTCCAAATGGATGGCTCTCTGAAACCTCACCCAGTGAAGCC 370

Qy 2922 CAGCTTTATTGGGACCCAGGCTTTTGAAGACTATGACCTGCAGAACTGGTGGACTACAT 2981
Db 371 CAGCTTTATTGGGACCCAGGCTTTTGAAGACTATGACCTGCAGAACTGGTGGACTACAT 430

Qy 2982 TGACTGGAAGCCTTTCTTGTGATGTCCTGGAGCTCCGGGCAAGTACCAGATCGAGCTT 3041
Db 431 TGACTGGAAGCCTTTCTTGTGATGTCCTGGAGCTCCGGGCAAGTACCAGATCGAGCTT 490

RESULT 30
BP332632
LOCUS      BP332632      583 bp      mRNA      linear      EST 17-SEP-2004
DEFINITION BP332632 Sugano cDNA library, renal proximal tubule primary
epithelial cell Homo sapiens CDNA clone RPR09337, mRNA sequence.
ACCESSION  BP332632
VERSION    BP332632.1 GI:52262059
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE  1 (bases 1 to 583)
AUTHORS   Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE     Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT   Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES             source
    Location/Qualifiers
        1..583
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="RPR09337"
            /tissue_type="renal proximal tubule"
            /cell_type="primary epithelial cell"
            /clone_lib="Sugano cDNA library, renal proximal tubule
primary epithelial cell"

ORIGIN
Query Match      6.4%; Score 251; DB 3; Length 593;
Best Local Similarity 100.0%; Pred. No. 6.5e-114;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2 GTGACCTGTGAGAGCAGCTTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 61
Db 333 GTGACCTGTGAGAGCAGCTTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 392
QY 62 ACATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGCTCTGAAGAAACCTCGCGG 121
Db 393 ACATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGCTCTGAAGAAACCTCGCGG 452
QY 122 ATGAGATCAATGCGCATTTCTGAGAAAGAGATTTATGTTGCGATGAGGAGATGGGACCA 181
Db 453 ATGAGATCAATGCGCATTTCTGAGAAAGAGATTTATGTTGCGATGAGGAGATGGGACCA 512
QY 182 TGATCCAGCGGGAGAGCTTAACCAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241
Db 513 TGATCCAGCGGGAGAGCTTAACCAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 572
QY 242 CCAGCGCGCTG 252
Db 573 CCAGCGCGCTG 583

RESULT 31
LOCUS BX484356
DEFINITION BX484356 466 bp mRNA linear EST 04-SEP-2003
ACCESSION DXFZp686H23242_r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
VERSION BX484356
KEYWORDS DXFZp686H23242 5', mRNA sequence.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 466)
AUTHORS Ootenwaelder,B., Obermaier,B., Deutschenbaur,S., Mewes,H.W.,
TITLE Well,B., Amd,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
JOURNAL EST (Ootenwaelder,B., Obermaier,B., Deutschenbaur,S., Mewes,H.W.,
COMMENT et al.)
Unpublished (2003)
CONTACT: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. sl sequence
also available.
This clone (DKFZp686H23242) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source Location/Qualifiers
1..466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DXFZp686H23242"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 6.0%; Score 236; DB 5; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.3e-106;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 TATGCTCTCTGTTATCCCAATGCGAGGCTTCCCAACACCTTTGGTGACTATGATGAACG 963
Db 56 TATGCTCTCTGTTATCCCAATGCGAGGCTTCCCAACACCTTTGGTGACTATGATGAACG 115

```

```

QY 964 CTTTCTATGATGCGCAAGCACCTAAAGGATTTTCTGCTATGATGCTTGGTCAATATAGTT 1023
Db 116 CTTTCTATGATGCGCAAGCACCTAAAGGATTTTCTGCTATGATGCTTGGTCAATATAGTT 175
QY 1024 GGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAATTCCTGAAGCTGTGAATAAT 1083
Db 176 GGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAATTCCTGAAGCTGTGAATAAT 235
QY 1084 TGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGG 1139
Db 236 TGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGG 291

RESULT 32
LOCUS BP348983
DEFINITION BP348983 Sugano cDNA library, brain Homo sapiens cDNA clone
ACCESSION SZR02673, mRNA sequence.
VERSION BP348983
KEYWORDS BP348983.1 GI:52278968
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
TITLE Mizushima-Sugano,J., Nakai,K. and Sugano,S.
JOURNAL Sequence comparison of human and mouse genes reveals a homologous
PUBMED block structure in the promoter regions
COMMENT Genome Res. 14 (9), 1711-1718 (2004)
15342556
CONTACT: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuki@ims.u-tokyo.ac.jp.
FEATURES
source Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SZR02673"
/tissue_type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 5.9%; Score 232; DB 3; Length 581;
Best Local Similarity 100.0%; Pred. No. 2.3e-104;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 61
Db 350 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 409
QY 62 ACATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGCTCTGAAGAAACCTCGCGG 121
Db 410 ACATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGCTCTGAAGAAACCTCGCGG 469
QY 122 ATGAGATCAATGCGCATTTCTGAGAAAGAGATTTATGTTGCGATGAGGAGATGGGACCA 181
Db 470 ATGAGATCAATGCGCATTTCTGAGAAAGAGATTTATGTTGCGATGAGGAGATGGGACCA 529
QY 182 TGATCCAGCGGGAGAGCTTAACCAAGAACACTTCCGAGGTCAGGAATTTAA 233
Db 530 TGATCCAGCGGGAGAGCTTAACCAAGAACACTTCCGAGGTCAGGAATTTAA 581

RESULT 33
LOCUS BP249465

```

BP249465 592 bp mRNA linear EST 15-SEP-2004

DEFINITION BP249465 Sugano cDNA library, embryo kidney Homo sapiens cDNA clone  
HKR13611, mRNA sequence.

ACCESSION BP249465

VERSION BP249465.1 GI:52131744

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 592)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

PUBMED 15342556

COMMENT

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. 592

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="HKR13611"

/tissue\_type="kidney"

/cell\_lines="293"

/dev\_stages="embryo"

/clone\_lib="Sugano cDNA library, embryo kidney"

ORIGIN

Query Match 5.9%; Score 231; DB 3; Length 592;

Best Local Similarity 100.0%; Pred. No. 7.5e-104;

Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCACCTGTGAGAGCAGCTTCTCTCGCGGCCCTCTGGCAAGGAGGAGCTCGACA 61

Db 362 GTCACCTGTGAGAGCAGCTTCTCTCGCGGCCCTCTGGCAAGGAGGAGCTCGACA 421

QY 62 ACATGTCAACCGCGCTCCAAAGACCTGTCGCAACCGAAGGCTGAAGAAAACCTTCGGG 121

Db 422 ACATGTCAACCGCGCTCCAAAGACCTGTCGCAACCGAAGGCTGAAGAAAACCTTCGGG 481

QY 122 ATGAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTCGATGGAGGATGGGACCA 181

Db 482 ATGAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTCGATGGAGGATGGGACCA 541

QY 182 TGATCCAGCGGAGAGCTTAACGAAACACTTCCGAGGTCAGGAATTTA 232

Db 542 TGATCCAGCGGAGAGCTTAACGAAACACTTCCGAGGTCAGGAATTTA 592

RESULT 34

BU429497

LOCUS

DEFINITION UI-HF-BN0-aem-h-04-0-UI.r1 NIH MGC\_50 Homo sapiens cDNA clone

IMAGE:3064927 5', mRNA sequence.

ACCESSION BU429497

VERSION BU429497.1 GI:22767984

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 496)

AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rc@mail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M.B. Soares Lab

cDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio-llnl.gov/bbrp/image/image.html](http://www-bio-llnl.gov/bbrp/image/image.html)

Seq primer: M13 Forward

Location/Qualifiers

1. 496

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3064927"

/tissue\_type="lymph"

/cell\_type="germinal center B cells"

/cell\_line="MGC85"

/lab\_host="DH10B (LTI)"

/clone\_lib="NIH\_MGC\_50"

/notes="Vector: pT73-Pac; Site 1: NotI; Site 2: Eco RI;

Constructed from size fractionated cytoplasmic mRNA

(3.5-4.4kb). Directionally cloned. Cells provided by

Louis M. Staudt, Ph.D. Library preparation by Maria de

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 5.5%; Score 216; DB 5; Length 496;

Best Local Similarity 99.2%; Pred. No. 2.6e-96;

Matches 366; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3165 CTGGCCAGCAGAGTATCCAGACGACATTCACCTGTACGACAGGCTGTGTGCCCA 3224

Db 8 CTGGCCAGCAGAGTATCCAGACGACATTCACCTGTACGACAGGCTGTGTGCCCA 67

QY 3225 GGCTGACAGAGCCATAGCCACTTTCTATGGTTAAGCAACAGAGCTGAGAGGACTCTGC 3284

Db 68 GGCTGACAGAGCCATAGCCACTTTCTATGGTTAAGCAACAGAGCTGAGAGGACTCTGC 127

QY 3285 CAGCACGAGGCATATCTACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCG 3344

Db 128 CAGCACGAGGCATATCTACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCG 187

QY 3345 TGACTACTGGGCTGTTGGCGTTGGCTGCTTTGGGGTAGAGAGCTGAGAGGAGGCTA 3404

Db 188 TGACTACTGGGCTGTTGGCGTTGGCTGCTTTGGGGTAGAGAGCTGAGAGGAGGCTA 247

QY 3405 TGAGGATGATGGTGACGACTACAGCAGCATCATGCTCAAGGCGCTGGGGGACCGGCTGGC 3464

Db 248 TGAGGATGATGGTGACGACTACAGCAGCATCATGCTCAAGGCGCTGGGGGACCGGCTGGC 307

QY 3465 AGAGGCTTTGACAGAGAGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCTTACTGTGG 3524

Db 308 AAAGGCTTTGACAGAGAGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCTTACTGTGG 367

QY 3525 CAGTGAGCA 3533

Db 368 CAGTGAGCA 376

RESULT 35

CV356711

LOCUS

DEFINITION MR4-RT0026-050201-301-h10 RT0026 Homo sapiens cDNA, mRNA sequence.

ACCESSION CV356711

VERSION CV356711.1 GI:52706766

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.





```

DKFZp781f1761 5', mRNA sequence.
ACCESSION BX955050
VERSION BX955050.1 GI:43436899
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 356)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
FOBO,G., Han,M. and Wiemann,S.
TITLE EST (Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No si sequence available.
This clone (DKFZp781f1761) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
Location/Qualifiers
1..356
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781f1761"
/dev_stage="adult"
/lab_host="DH10B"
/clone_libs="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIB;
cDNA-collection"
ORIGIN
Query Match 5.1%; Score 201; DB 5; Length 356;
Best Local Similarity 100.0%; Pred. No. 8.9e-89;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 518 CACTCTCTGTGCCCCATCTGTGGAAGCGCGGATTATAGGAACATCATCTTGATGAGC 577
Db 201 CACTCTCTGTGCCCCATCTGTGGAAGCGCGGATTATAGGAACATCATCTTGATGAGC 142
Qy 578 TTGTTGAAGCATACCAAGAGCAGGCGCAAGGACCTTCTGGATGGCGGGGTTGATATCTTAC 637
Db 141 TTGTTGAAGCATACCAAGAGCAGGCGCAAGGACCTTCTGGATGGCGGGGTTGATATCTTAC 82
Qy 638 TCATTGAAACTATTTTGTATCTGCAATGCCAAGGAGCCCTGTTGTCACCTCCAAATC 697
Db 81 TCATTGAAACTATTTTGTATCTGCAATGCCAAGGAGCCCTGTTGTCACCTCCAAATC 22
Qy 698 TTTTGGAGGAAATATGCTC 718
Db 21 TTTTGGAGGAAATATGCTC 1
RESULT 38
BP248483 Sugano cDNA library, embryo kidney Homo sapiens cDNA clone
LOCUS HKR111133, mRNA sequence.
DEFINITION BP248483
ACCESSION BP248483
VERSION BP248483.1 GI:52130762
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 996)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapps-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13532 row: i column: 07
High quality sequence stop: 181.

REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezukie@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HKR111133"
/tissue_type="kidney"
/cell_line="293"
/dev_stage="embryo"
/clone_lib="Sugano cDNA library, embryo kidney"
ORIGIN
Query Match 4.6%; Score 180; DB 3; Length 581;
Best Local Similarity 99.6%; Pred. No. 3.3e-78;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGAGAGAGAGCTCGACA 61
Db 351 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGAGAGAGAGCTCGACA 410
Qy 62 ACATGTCAACCGCGCTCCAGAGCTGTGCGAACCAGAGCTCTGAGAAACCTCGCGGG 121
Db 411 ACATGTCAACCGCGCTCCAGAGCTGTGCGAACCAGAGCTCTGAGAAACCTCGCGGG 470
Qy 122 ATGAGATCAATGCCATTCTGCGAAGAGAGATTATGCTGCTGGATGGAGGATGGGACCA 181
Db 471 ATGAGATCAATGCCATTCTGCGAAGAGAGATTATGCTGCTGGATGGAGGATGGGACCA 530
Qy 182 TGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTA 232
Db 531 TGATCCAGCGGAGAGCTAAACGAAGAACACTTCCGAGGTCAGGAATTTA 581
RESULT 39
BU184298 996 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT 7858169 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168654
DEFINITION 5', mRNA sequence.
ACCESSION BU184298
VERSION BU184298.1 GI:22698282
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 996)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapps-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13532 row: i column: 07
High quality sequence stop: 181.

```

```
FEATURES
  source
    Location/Qualifiers
      1..996
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6168654"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH MGC 72"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."
      ORIGIN
        Query Match 4.0%; Score 158; DB 5; Length 996;
        Best Local Similarity 99.5%; Pred. No. 3.9e-67;
        Matches 208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
      QY 2675 AAATAGTCCGAGATACAGTGCACCTGTATCCATGCTCTGGACGCGTCCAAAGTGTGG 2734
      Db 1 AAATAGTCCGAGATACAGTGCACCTGTATCCATGCTCTGGACGCGTCCAAAGTGTGG 60
      QY 2735 TGGTGTCTCCAGCTGTAGATGAATCTAAAGGATGAATCTTTGAGGAAATCATGG 2794
      Db 61 TGGTGTCTCCAGCTGTAGATGAATCTAAAGGATGAATCTTTGAGGAAATCATGG 120
      QY 2795 AAGAATATGAATATTAGACAGGACCATTTATGAGTCTCTCAAGGAGGAGATCTTAC 2854
      Db 121 AAGAATATGAATATTAGACAGGACCATTTATGAGTCTCTCAAGGAGGAGATCTTAC 180
      QY 2855 CCTTAAGTCAAGCCAGAAAAGTGGTTTC 2883
      Db 181 CCTTAAGTCAAGCCAGAAAAGTGGTTTC 209
    RESULT 40
    A1872333/c
    LOCUS
    DEFINITION
      wh57h11.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2440101 3'
      similar to SW:METH HUMAN Q99707
      5-METHYLTRANSFERASE--HOMOCYSTEINE METHYLTRANSFERASE ; mRNA
      sequence.
    ACCESSION
      A1872333
    VERSION
      A1872333.1 GI:5546382
    KEYWORDS
      EST.
    SOURCE
      Homo sapiens (human)
    ORGANISM
      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Homnidae; Homo.
    REFERENCE
      1 (bases 1 to 668)
      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
      Tumor Gene Index
      Unpublished (1997)
      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs@mail.nih.gov
      Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
      Emmert-Buck, M.D., Ph.D.
      cDNA Library Preparation: Life Technologies, Inc.
      cDNA Library Arrayed by: Greg Lemmon, Ph.D.
      DNA Sequencing by: Washington University Genome Sequencing Center
      Clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      www-bio.llnl.gov/bbrp/image/image.html
      Insert Length: 2769 Std Error: 0.00
      Seq primer: -40UP from Gibco
      High quality sequence stop: 404.
      Location/Qualifiers
        1..668
        /organism="Homo sapiens"
        /mol_type="mRNA"
      FEATURES
        source
          /db_xref="taxon:9606"
          /clone="IMAGE:2440101"
          /tissue_type="moderately-differentiated endometrial
          adenocarcinoma, 3 pooled tumors"
          /lab_host="DH10B"
          /clone_lib="NCI CGAP Ut2"
          /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
          Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
          Average insert size 1.85 Kb. Life Technologies catalog #:
          11539-012"
        ORIGIN
          Query Match 3.9%; Score 154; DB 1; Length 668;
          Best Local Similarity 100.0%; Pred. No. 4e-65;
          Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
        QY 3724 TCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCGAAGGATCAGGTTGAGGAT 3783
        Db 252 TCCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCGAAGGATCAGGTTGAGGAT 193
        QY 3784 TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCATTTTG 3843
        Db 192 TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCATTTTG 133
        QY 3844 GGATATGATACACACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3877
        Db 132 GGATATGATACACACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 99
    RESULT 41
    CV363529
    LOCUS
    DEFINITION
      PM1-CN0098-110101-009-a05 CN0098 Homo sapiens cDNA, mRNA sequence.
    ACCESSION
      CV363529
    VERSION
      CV363529.1 GI:52713584
    KEYWORDS
      EST.
    SOURCE
      Homo sapiens (human)
    ORGANISM
      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Homnidae; Homo.
    REFERENCE
      1 (bases 1 to 477)
      Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
      Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
      Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
      Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
      O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
      Simpson, A.J.
      Shotgun sequencing of the human transcriptome with ORF expressed
      sequence tags
      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
      10737800
    COMMENT
      Contact: Simpson A.J.G.
      Laboratory of Cancer Genetics
      Ludwig Institute for Cancer Research
      Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
      Brazil
      Tel: +55-11-2704922
      Fax: +55-11-2707001
      Email: asimpson@ludwig.org.br
      This sequence was derived from the FAPESP/LICR Human Cancer Genome
      Project. http://www.ludwig.org.br.
      Location/Qualifiers
        1..477
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="CN0098"
        /note="Organ: colon normal; Vector: puc18; Site 1: SmaI;
        Site 2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
```

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 3.6%; Score 142; DB 7; Length 477;  
Best Local Similarity 99.5%; Pred. No. 4.2e-59;  
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 3366 CGTTGCTGCTTTGGGTAGAGAGCTGAGCAAGCCCTATGAGATGATGTCGAGCTA 3425  
Db 152 CGTTGCTGCTTTGGGTAGAGAGCTGAGCAAGCCCTATGAGATGATGTCGAGCTA 211  
  
Qy 3426 CAGCAGCATCATGCTCAAGGCGCTGGGGAGCCGCTGCAGAGCCCTTTGCAGAGAGCT 3485  
Db 212 CGGCAGCATCATGCTCAAGGCGCTGGGGAGCCGCTGCAGAGCCCTTTGCAGAGAGCT 271  
  
Qy 3486 CCATGAAGAGTGTGCCGAGAACTGTGGCCCTACTGTGGCAGTGAGCAGCTGGAGCTGC 3545  
Db 272 CCATGAAGAGTGTGCCGAGAACTGTGGCCCTACTGTGGCAGTGAGCAGCTGGAGCTGC 331  
  
Qy 3546 AGACTGCGAAGG 3558  
Db 332 AGACTGCGAAGG 344

## RESULT 42

LOCUS CR767594 552 bp mRNA linear EST 23-SEP-2004  
DEFINITION DKFZp4690148 r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone  
ACCESSION DKFZp4690148.5, mRNA sequence.  
VERSION CR767594.1 GI:52609891  
KEYWORDS EST.  
SOURCE Pongo pygmaeus (orangutan)  
ORGANISM Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pongo.  
1 (bases 1 to 552)  
REFERENCE Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,  
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and  
Wiemann,S.  
Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,  
Deutschenbaur,S., et al.)  
JOURNAL Unpublished (2004)  
COMMENT Contact: MIPS  
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
Email s.wiemann@kfz-heidelberg.de; sequenced by Medigenomix  
(Martinried/Germany) within the cDNA sequencing consortium of the  
German Genome Project. This clone (DKFZp4690148) is available at  
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in  
Berlin, Germany. Please contact RZPD for ordering.  
http://www.rzpd.de/cgi-bin/products/ci.cgi?CloneId=DKFZp4690148  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/.

## FEATURES

source  
1..552  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp4690148"  
/tissue\_type="kidney"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="469 (synonym: pkid1)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

## ORIGIN

Query Match 3.5%; Score 136; DB 7; Length 552;

Best Local Similarity 99.2%; Pred. No. 4.4e-56;  
Matches 236; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1391 AGGTACCTTTGTGATCGACTCTCCAAATTTGCTGTGATTGAAGCTGGGTAAAGTGCT 1450  
Db 95 AGGTACCTTTGTGATCGACTCTCCAAATTTGCTGTGATTGAAGCTGGGTAAAGTGCT 154  
  
Qy 1451 GCCAAGGGAAAGTCATTGTCAATAGCATTAGTCTGAAGGAAAGAGAGAGACGACTTCTGG 1510  
Db 155 GCCAAGGGAAAGTCATTGTCAATAGCATTAGTCTGAAGGAAAGAGAGAGACGACTTCTGG 214  
  
Qy 1511 AGAAGGCCAGCAAGATTAAAAAGTATCGAGCTGCTATGGTGGTCATGGCTTTTGATGAAG 1570  
Db 215 AGAAGGCCAGCAAGATTAAAAAGTATCGAGCTGCTGCTGGTGGTTCATGGCTTTTGATGAAG 274  
  
Qy 1571 AAGCAGAGGCACAGAAACAGACACAAAATTCAGAGTGTGCACCCGGCCCTACCATCT 1628  
Db 275 AAGCAGAGGCACAGAAACAGACACAAAATTCAGAGTGTGCACCCGGCCCTACCATCT 332

## RESULT 43

LOCUS BF749082/c 226 bp mRNA linear EST 10-JAN-2001  
DEFINITION MR2-BN0386-051000-014-f10 BN0386 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF749082  
VERSION BF749082.1 GI:12075758  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 226)  
REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-BN0386-  
051000-014-f10&t3=2000-10-05&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 226.  
Location/Qualifiers  
1..226  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BN0386"

## FEATURES

source  
1..226  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BN0386"  
/note="Organ: breast normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORSTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN



**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Homo.

**REFERENCE** 1 (bases 1 to 1168)  
**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgsbbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2645 row: a column: 19  
 High quality sequence stop: 131.

**FEATURES** Location/Qualifiers  
 1..1168  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6459642"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 101"  
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

**ORIGIN**  
 Query Match 3.2%; Score 127; DB 5; Length 1168;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-51;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3361 TTTCGGTGGCTGCTTTGGGTAGACAGCTGACAGGCGCTATGAGGATGATGGTGAC 3420  
 |||||  
 Db 1 TTTCGGTGGCTGCTTTGGGTAGACAGCTGACAGGCGCTATGAGGATGATGGTGAC 60  
 |||||

QY 3421 GACTACAGCAGCATCGTCAAGCGCTGGGGACCGCTGGCAGAGCGCTTTGCAGAA 3480  
 |||||  
 Db 61 GACTACAGCAGCATCGTCAAGCGCTGGGGACCGCTGGCAGAGCGCTTTGCAGAA 120  
 |||||

QY 3481 GAGCTCC 3487  
 |||||  
 Db 121 GAGCTCC 127  
 |||||

**RESULT 47**  
**CR853354**  
**LOCUS** CR853354 269 bp mRNA linear EST 22-OCT-2004  
**DEFINITION** DKFZp469J177 r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone  
 DKFZp469J177 5', mRNA sequence.  
**ACCESSION** CR853354  
**VERSION** CR853354.1 GI:54568166  
**KEYWORDS** EST.  
**SOURCE** Pongo pygmaeus (orangutan)  
**ORGANISM** Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Pongo.

**REFERENCE** 1 (bases 1 to 269)  
**AUTHORS** Ottenwaelder,B.; Obermaier,B., Deutschenbaur,S., Schapp,A.,  
 Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and  
 Wiemann,S.

**TITLE** Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,  
 Deutschenbaur,S., et al.)  
**JOURNAL** Unpublished (2004)  
**COMMENT** Contact: MIPS  
 MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix  
 (Martinsried/Germany) within the cDNA sequencing consortium of the  
 German Genome Project. This clone (DKFZp469J177) is available at  
 the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in  
 Berlin, Germany. Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469J177>  
 Further information about the clone and the sequencing project is  
 available at <http://mips.gsf.de/projects/cdna/>.

**FEATURES** Location/Qualifiers  
 1..269  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp469J177"  
 /tissue\_type="kidney"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="469 (synonym: pkid1)"  
 /note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

**ORIGIN**  
 Query Match 3.2%; Score 124; DB 7; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-50;  
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1391 AGGTACCTTTGTGATCGATCTCTCCAAATTTGCTGTGATGAGCTGGGTTAAAGTGCT 1450  
 |||||  
 Db 146 AGGTACCTTTGTGATCGATCTCTCCAAATTTGCTGTGATGAGCTGGGTTAAAGTGCT 205  
 |||||

QY 1451 GCCAAGGGAGTGCATTGTCAATAGCATTAGTCTGAAGGAAGGAGACGACTTCTTGG 1510  
 |||||  
 Db 206 GCCAAGGGAGTGCATTGTCAATAGCATTAGTCTGAAGGAAGGAGACGACTTCTTGG 265  
 |||||

QY 1511 AGAA 1514  
 |||||  
 Db 266 AGAA 269  
 |||||

**RESULT 48**  
**AQ744889**  
**LOCUS** AQ744889 857 bp DNA linear GSS 16-JUL-1999  
**DEFINITION** HS 5507\_Al\_D02.SP6.RPCI-11 Human Male BAC Library Homo sapiens  
 genomic clone Plate=1083 Col=3 Row=G, genomic survey sequence.  
**ACCESSION** AQ744889  
**VERSION** AQ744889.1 GI:5522411  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Homo.  
**REFERENCE** 1 (bases 1 to 857)  
**AUTHORS** Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
 Hood,L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 10449764  
**TITLE** Contact: Mahairas GG, Wallace JC, Hood L  
**JOURNAL** High Throughput Sequencing Center  
**PUBMED** University of Washington  
**COMMENT** 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887



Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ([http://bacpac.med.buffalo.edu/ordering\\_bac.htm](http://bacpac.med.buffalo.edu/ordering_bac.htm)) or from Research Genetics (<http://www.resgen.com>). BAC end Web Server: <http://www.htsc.washington.edu>  
 Plate: 1083 row: G column: 3  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 857.  
 Location/Qualifiers  
 1..857  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate:1083 Col=3 Row=G"  
 /sex="male"  
 /clone\_lib="RPCI-11 Human Male BAC Library"  
 /note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

# FEATURES

source

# ORIGIN

Query Match 3.0%; Score 117; DB 9; Length 857;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-46;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1759 GAAACATTACCTGAGCCAGAAATAGTGGAGGTCTTTCCAACTTGCTTCTCTCCGA 1818  
 Db 281 GAAACATTACCTGAGCCAGAAATAGTGGAGGTCTTTCCAACTTGCTTCTCTCCGA 340  
 Qy 1819 GGAATGGAAGCCATTGAGAGAGCAATGCATCGGGTTTTCCTTACCATGCAATCAAG 1875  
 Db 341 GGAATGGAAGCCATTGAGAGAGCAATGCATCGGGTTTTCCTTACCATGCAATCAAG 397

RESULT 49  
 AG092782  
 LOCUS AG092782 735 bp DNA linear GSS 03-NOV-2001  
 DEFINITION Pan troglodytes DNA, clone: PTB-093A19.R, genomic survey sequence.  
 ACCESSION AG092782  
 VERSION AG092782.1 GI:16644584  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1  
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE BAC end sequences of Library PTB  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 735)  
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbes@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
 COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  
 PRIMERS  
 Sequencing: M13Rev  
 LIBRARY  
 Vector : pKS145  
 R.Site 1 : SacI

R.Site 2 : SacI.  
 Location/Qualifiers  
 1..735  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-093A19.R"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"

# ORIGIN

Query Match 2.9%; Score 114; DB 10; Length 735;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-45;  
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2255 CTCAGGTATATAAGTCAGCCGGTATGAAGAAGGCTGTGGCCACTTATCCCTTTCA 2314  
 Db 452 CTCAGGTATATAAGTCAGCCGGTATGAAGAAGGCTGTGGCCACTTATCCCTTTCA 511  
 Qy 2315 TGGAAAAGAAAGAGAGAAACACAGAGTGTCTTAACCGCACAGTAGAAGAGAGG 2368  
 Db 512 TGGAAAAGAAAGAGAGAAACACAGAGTGTCTTAACCGCACAGTAGAAGAGAGG 565

# RESULT 50

CF135946  
 LOCUS CF135946 153 bp mRNA linear EST 09-SEP-2003  
 DEFINITION UT-HB-BNO-amc-c-02-0-UI.r1 NIH\_MGC\_50 Homo sapiens cDNA clone IMAGE:3090794 5', mRNA sequence.  
 ACCESSION CF135946  
 VERSION CF135946.1 GI:33251390  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 153)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Louis Staudt  
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/humanfl.html>  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1..153  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3090794"  
 /tissue\_type="lymph"  
 /cell\_type="germinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DH10B (LTI)"  
 /clone\_lib="NIH\_MGC\_50"  
 /note="Vector: pT713-Pac; Site 1: NotI; Site 2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de

# FEATURES

source

Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D. "

## ORIGIN

Query Match 2.5%; Score 98; DB 6; Length 153;  
Best Local Similarity 99.3%; Pred. No. 5.4e-37;  
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3398 AGGCTATGAGGATGATGGTGACGACTACAGCAGCATCATGGTCAAGCGCTGGGGACC 3457  
|||  
Db 5 AGGCTATGAGGATGATGGTGACGACTACAGCAGCATCATGGTCAAGCGCTGGGGACC 64  
|||

Qy 3458 GGCTGGCAGAGGCCCTTTGACAGAGAGCTCCATGAAAGATTCCCGAGAACTGTGGGCCT 3517  
|||  
Db 65 GGCTGGCANAGGCCCTTTGACAGAGAGCTCCATGAAAGATTCCCGAGAACTGTGGGCCT 124  
|||

Qy 3518 ACTGTGGCAGTGAGCAGCTGGACGTCCGA 3546  
|||  
Db 125 ACTGTGGCAGTGAGCAGCTGGACGTCCGA 153  
|||

Search completed: March 7, 2006, 10:54:49  
Job time : 13647 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 01:43:57 ; Search time 18876 Seconds  
(without alignments)  
11801.732 Million cell updates/sec

Title: US-10-607-712-1  
Perfect score: 3919  
Sequence: 1 ggtcactgtgagacg.....ctcaaggaaatacaacctag 3919

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 18 *or more - this run search includes oligomers 18-60.*  
Total number of hits satisfying chosen parameters: 196778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_sv.\*
- 12: gb\_un.\*
- 13: gb\_vl.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3919	100.0	3919	6 AR482278	AR482278 Sequence
2	3823	97.6	3917	8 HSU71285	U71285 Human 5-met
3	3601	91.9	3856	6 AR482320	AR482320 Sequence
4	3517	89.7	7122	6 AR144957	AR144957 Sequence
5	3517	89.7	7122	6 AR367907	AR367907 Sequence
6	3517	89.7	7122	6 AX050442	AX050442 Sequence
7	3517	89.7	7122	6 AX069340	AX069340 Sequence
8	3517	89.7	7122	8 HSU75743	U75743 Human methi
9	3466	88.4	7224	6 AR367908	AR367908 Sequence
10	3466	88.4	7224	6 AX069342	AX069342 Sequence
11	3466	88.4	7224	8 HSU73338	U73338 Human methi
12	3390	86.5	7398	6 AR930571	AR930571 Sequence
13	3109	79.3	7224	6 AR300095	AR300095 Sequence
14	3109	79.3	7224	6 AR438495	AR438495 Sequence
15	3109	79.3	7224	6 AR562003	AR562003 Sequence
16	1779	45.4	5240	8 HSM807075	EX641132 Homo sapi
17	714	18.2	1835	6 CQ721809	CQ721809 Sequence
18	308	7.9	4172	6 CQ850206	CQ850206 Sequence

19	308	7.9	4172	8	AK127317	Homo sapi
20	297	7.6	435	8	AY429553	Homo sapi
21	282	5.7	305	6	BD060282	Secreted
22	217	5.5	108841	8	AL359185	Human DNA
23	217	5.5	183166	14	AL357554	Homo sapi
24	204	5.2	164529	14	AC079975	Homo sapi
25	204	5.2	164933	14	AC074231	Homo sapi
26	204	5.2	172753	8	AL359259	Human DNA
27	204	5.2	176671	14	AC080149	Homo sapi
28	189	4.8	183166	14	AL357554	Homo sapi
29	183	4.7	176671	14	AC080149	Homo sapi
30	182	4.6	73656	14	AC016512	Homo sapi
31	129	3.3	201	10	BV204261	sgnm21456
32	129	3.3	67611	14	AC080113	Homo sapi
33	120	3.1	67611	14	AC080113	Homo sapi
34	114	2.9	164529	14	AC079975	Homo sapi
35	103	2.6	73656	14	AC016512	Homo sapi
36	96	2.4	732	6	CQ750107	Sequence
37	95	2.4	731	10	BV599890	S217P6097
38	93	2.4	758	10	BV597233	S216P6112
39	68	1.7	3832	4	DQ084519	Bos tauru
40	65	1.7	4498	9	AF034214	Rattus no
41	63	1.6	63	6	AX611831	Sequence
42	57	1.5	211299	14	AC156661	Bos tauru
43	55	1.4	55	6	AX611827	Sequence
44	54	1.4	55	6	AX611829	Sequence
45	54	1.4	194232	14	AC154523	Mus muscu
46	54	1.4	228824	14	AC137163	Rattus no
47	54	1.4	236536	9	AC139941	Mus muscu
48	54	1.4	244504	14	AC113729	Rattus no
49	54	1.4	268715	14	AC121488	Rattus no
50	41	1.0	5168	5	AJ720857	Gallus ga
51	38	1.0	402	4	AY128710	Bos tauru
52	38	1.0	489	4	AF276463	Sus scro
53	37	0.9	63	6	AX611832	Sequence
54	36	0.9	55	6	AX611830	Sequence
55	34	0.9	55	6	AX611828	Sequence
56	28	0.7	28	6	AR482298	Sequence
57	28	0.7	54625	14	AC137880	Mus muscu
58	27	0.7	27	6	CQ785883	Sequence
59	27	0.7	27	6	AX614442	Sequence
60	27	0.7	144781	8	HSJ585114	Human DNA
61	27	0.7	151843	14	AC155050	Bos tauru
62	27	0.7	165307	9	AC132116	Mus muscu
63	27	0.7	181349	9	AL772214	Mouse DNA
64	27	0.7	185332	8	AC016673	Homo sapi
65	27	0.7	186566	5	BX537335	Zebrafish
66	27	0.7	199128	14	CR759764	Danio rer
67	27	0.7	250135	14	AC161685	Bos tauru
68	26	0.7	26	6	AR482288	Sequence
69	26	0.7	26	6	AR482301	Sequence
70	26	0.7	26	6	AR482306	Sequence
71	26	0.7	26	6	AR482310	Sequence
72	26	0.7	26	6	AR482316	Sequence
73	26	0.7	995	5	BC075365	Xenopus t
74	26	0.7	2051	4	CFGFCR1	Canis fami
75	26	0.7	7387	8	AC109519	Homo sapi
76	26	0.7	37495	14	AC017418	Drosophila
77	26	0.7	64848	8	AL603746	Human DNA
78	26	0.7	69026	14	AC105115	Homo sapi
79	26	0.7	80919	2	AC092191	Drosophila
80	26	0.7	82711	8	AC103916	Homo sapi
81	26	0.7	85346	14	AC165617	Bos tauru
82	26	0.7	90175	8	AC092391	Homo sapi
83	26	0.7	112020	14	AC133285	Homo sapi
84	26	0.7	113041	8	AL391297	Human DNA
85	26	0.7	148496	14	AC079332	Homo sapi
86	26	0.7	153920	14	AC011723	Homo sapi
87	26	0.7	154749	4	AB113354	Sus scro
88	26	0.7	158432	14	AC105438	Sus scro
89	26	0.7	164287	9	AL691422	Mouse DNA
90	26	0.7	169246	14	AC135840	Sus scro
91	26	0.7	171329	4	CR956646	Pig DNA B

C 92	26	0.7	172266	2	AC092190	AC092190 Drosophila	165	25	0.6	141656	14	AC163075	AC163075 Bos taurus
C 93	26	0.7	172638	14	AC155863	AC155863 Bos taurus	166	25	0.6	142976	6	CR0869704	CR0869704 Sequence
C 94	26	0.7	174874	14	AC026259	AC026259 Homo sapi	167	25	0.6	143454	14	AC152489	AC152489 Dasyatis n
C 95	26	0.7	175081	8	AC021151	AC021151 Homo sapi	168	25	0.6	145231	9	AC119869	AC119869 Mus muscu
C 96	26	0.7	178594	14	AC139455	AC139455 Homo sapi	169	25	0.6	147032	5	CR759945	CR759945 Zebrafish
C 97	26	0.7	179691	4	AC139455	AC139455 Homo sapi	170	25	0.6	147350	14	AC069332	AC069332 Homo sapi
C 98	26	0.7	181256	14	AC150611	AC150611 Callithrix	171	25	0.6	151160	14	AC151022	AC151022 Callithrix
C 99	26	0.7	181432	14	AC134962	AC134962 Sus scrofa	172	25	0.6	152244	8	AP000352	AP000352 Homo sapi
C 100	26	0.7	181774	14	AC159637	AC159637 Mus muscu	173	25	0.6	153469	14	CR927192	CR927192 Danio rer
C 101	26	0.7	184056	14	AC141857	AC141857 Sus scrofa	174	25	0.6	154101	14	AC136885	AC136885 Sus scrofa
C 102	26	0.7	190275	14	AC138042	AC138042 Sus scrofa	175	25	0.6	155724	8	AL138781	AL138781 Human DNA
C 103	26	0.7	191122	14	AL359769	AL359769 Homo sapi	176	25	0.6	155724	4	AC091316	AC091316 Sus scrofa
C 104	26	0.7	191381	14	AC138046	AC138046 Homo sapi	177	25	0.6	156721	8	AL139389	AL139389 Human DNA
C 105	26	0.7	205796	14	AL713891	AL713891 Mus muscu	178	25	0.6	157972	8	HSB782K7	HSB782K7 Human DNA
C 106	26	0.7	208079	9	AC153555	AC153555 Mus muscu	179	25	0.6	157974	14	AC121070	AC121070 Sus scrofa
C 107	26	0.7	215524	9	AL603913	AL603913 Mouse DNA	180	25	0.6	158432	14	AC105438	AC105438 Sus scrofa
C 108	26	0.7	217740	14	AC139949	AC139949 Rattus no	181	25	0.6	158845	14	AC150171	AC150171 Gallus ga
C 109	26	0.7	223708	9	AC133607	AC133607 Mus muscu	182	25	0.6	160659	14	AC036235	AC036235 Homo sapi
C 110	26	0.7	226568	9	AC148012	AC148012 Mus muscu	183	25	0.6	160854	8	AC110602	AC110602 Homo sapi
C 111	26	0.7	231946	9	AC153836	AC153836 Mus muscu	184	25	0.6	163787	4	AC087424	AC087424 Sus scrofa
C 112	26	0.7	236536	9	AC139941	AC139941 Mus muscu	185	25	0.6	164808	8	HSB7293L8	HSB7293L8 Human DNA
C 113	26	0.7	239539	14	AC162080	AC162080 Bos taurus	186	25	0.6	165704	14	CR956650	CR956650 Sus scrofa
C 114	26	0.7	241886	9	AC027285	AC027285 Mus muscu	187	25	0.6	166014	8	AC022538	AC022538 Homo sapi
C 115	26	0.7	243502	14	CR974422	CR974422 Mus muscu	188	25	0.6	166098	8	AC020641	AC020641 Homo sapi
C 116	26	0.7	245439	9	MMHC135G15	MMHC135G15 Mus muscu	189	25	0.6	166111	14	AC127749	AC127749 Rattus no
C 117	26	0.7	255393	14	CR974457	CR974457 Mus muscu	190	25	0.6	166835	8	AL133216	AL133216 Human DNA
C 118	26	0.7	259661	14	AC023911	AC023911 Homo sapi	191	25	0.6	168012	4	CR956391	CR956391 Pig DNA s
C 119	26	0.7	313035	2	AE003573	AE003573 Drosophila	192	25	0.6	168062	5	AC147788	AC147788 Latimeria
C 120	25	0.6	25	6	AR482292	AR482292 Sequence	193	25	0.6	168364	14	AC016066	AC016066 Homo sapi
C 121	25	0.6	25	6	AR482294	AR482294 Sequence	194	25	0.6	168609	14	AC140986	AC140986 Rattus no
C 122	25	0.6	25	6	AR482302	AR482302 Sequence	195	25	0.6	171251	14	AC123981	AC123981 Pan trogl
C 123	25	0.6	25	6	AR482311	AR482311 Sequence	196	25	0.6	172816	8	AC123981	AC123981 Pan trogl
C 124	25	0.6	25	6	AR482312	AR482312 Sequence	197	25	0.6	173804	4	AB210286	AB210286 Sus scrofa
C 125	25	0.6	25	6	AR482313	AR482313 Sequence	198	25	0.6	174443	14	AC025515	AC025515 Homo sapi
C 126	25	0.6	25	6	AR482314	AR482314 Sequence	199	25	0.6	174711	8	AC025515	AC025515 Homo sapi
C 127	25	0.6	25	6	AR482315	AR482315 Sequence	200	25	0.6	176085	8	AC021504	AC021504 Homo sapi
C 128	25	0.6	25	6	AR482318	AR482318 Sequence	201	25	0.6	177068	8	AC021504	AC021504 Homo sapi
C 129	25	0.6	237	6	CQ529205	CQ529205 Sequence	202	25	0.6	177352	14	AC146673	AC146673 Otlemur
C 130	25	0.6	296	6	CQ466754	CQ466754 Sequence	203	25	0.6	178315	8	AC122688	AC122688 Homo sapi
C 131	25	0.6	346	10	BV427957	BV427957 S237P631F	204	25	0.6	178413	14	AP001370	AP001370 Homo sapi
C 132	25	0.6	412	15	AF298347	AF298347 Gastrolob	205	25	0.6	182133	14	AC151102	AC151102 Bos taurus
C 133	25	0.6	742	10	BV595021	BV595021 S215P6155	206	25	0.6	182503	9	AL672195	AL672195 Mouse DNA
C 134	25	0.6	789	6	BD145240	BD145240 Primer fo	207	25	0.6	183149	14	AC130353	AC130353 Homo sapi
C 135	25	0.6	789	6	AX865178	AX865178 Sequence	208	25	0.6	184370	14	AC136887	AC136887 Sus scrofa
C 136	25	0.6	1180	6	BD231678	BD231678 31 human	209	25	0.6	184864	8	AC013553	AC013553 Homo sapi
C 137	25	0.6	1180	6	CQ972159	CQ972159 Sequence	210	25	0.6	185081	14	CT009693	CT009693 Sus scrofa
C 138	25	0.6	1724	4	AY332501S1	AY332501 Sus scrofa	211	25	0.6	185824	14	AC156317	AC156317 Bos taurus
C 139	25	0.6	1773	8	BC069185	BC069185 Homo sapi	212	25	0.6	187186	14	CR392353	CR392353 Panio rer
C 140	25	0.6	1838	5	BC074656	BC074656 Xenopus t	213	25	0.6	187377	5	CR933822	CR933822 Zebrafish
C 141	25	0.6	1895	4	AY332500S2	AY332500 Sus scrofa	214	25	0.6	187856	14	AC146933	AC146933 Sus scrofa
C 142	25	0.6	1974	8	BC050075	BC050075 Homo sapi	215	25	0.6	188517	9	AC136513	AC136513 Mus muscu
C 143	25	0.6	2045	5	AJ719756	AJ719756 Gallus ga	216	25	0.6	188853	9	AC102105	AC102105 Mus muscu
C 144	25	0.6	2049	8	BC028304	BC028304 Homo sapi	217	25	0.6	189091	9	AC112263	AC112263 Mus muscu
C 145	25	0.6	2705	9	BC050788	BC050788 Mus muscu	218	25	0.6	189985	14	AC140099	AC140099 Sus scrofa
C 146	25	0.6	7215	2	DMP13K68D	X92892 D.melanog	219	25	0.6	190275	14	AC138042	AC138042 Sus scrofa
C 147	25	0.6	14955	2	AF000580	AF000580 Dictyoste	220	25	0.6	192063	8	AC099665	AC099665 Homo sapi
C 148	25	0.6	65825	8	AL161904	AL161904 Human DNA	221	25	0.6	195342	8	AC009120	AC009120 Homo sapi
C 149	25	0.6	79594	8	AL138805	AL138805 Human DNA	222	25	0.6	196195	9	AC122383	AC122383 Mus muscu
C 150	25	0.6	79991	14	AC150181	AC150181 Gallus ga	223	25	0.6	196311	8	CNS01DVP	AL139999 Human chr
C 151	25	0.6	81878	14	AC013647	AC013647 Homo sapi	224	25	0.6	196357	14	AC163960	AC163960 Sus scrofa
C 152	25	0.6	83000	8	AP002960	AP002960 Homo sapi	225	25	0.6	196567	9	AL928855	AL928855 Mouse DNA
C 153	25	0.6	84444	4	CR974452	CR974452 Pig DNA s	226	25	0.6	196724	9	AL590963	AL590963 Mouse DNA
C 154	25	0.6	103310	14	AC163688	AC163688 Mus muscu	227	25	0.6	196782	14	AL390956	AL390956 Homo sapi
C 155	25	0.6	106008	8	HS06	AL020993 Human DNA	228	25	0.6	198515	14	RN323M7	CR937049 Rattus no
C 156	25	0.6	106730	15	ATF12M12	AL355775 Arabidops	229	25	0.6	198784	14	AC154888	AC154888 Bos taurus
C 157	25	0.6	110000	14	AC122076	AC122076 Rattus no	230	25	0.6	198935	14	AC068573	AC068573 Homo sapi
C 158	25	0.6	110000	14	AC156034 <sup>1</sup>	Continuation (2 of	231	25	0.6	200598	14	AC012505	AC012505 Homo sapi
C 159	25	0.6	110000	14	AC156034 <sup>2</sup>	Continuation (3 of	232	25	0.6	202377	9	AC109165	AC109165 Mus muscu
C 160	25	0.6	112732	14	AC092367	AC092367 Homo sapi	233	25	0.6	202381	14	AC166026	AC166026 Oryctolag
C 161	25	0.6	117614	14	AC151688	AC151688 Gallus ga	234	25	0.6	204138	4	SSC427478	AL724778 Sus scrofa
C 162	25	0.6	124055	14	AC127600	AC127600 Homo sapi	235	25	0.6	204177	9	AC124494	AC124494 Homo sapi
C 163	25	0.6	137083	14	AP007630	AP007630 Lotus cor	236	25	0.6	206119	14	AL932168	AL932168 Homo sapi
C 164	25	0.6	140367	9	AL732439	AL732439 Mouse DNA	237	25	0.6	206513	5	AL928728	AL928728 Zebrafish

C 238	25	0.6	206533	8	CNS01RH1	AL160237 Human chr	C 311	24	0.6	2009	9	BC027202	BC027202 Mus muscu
C 239	25	0.6	206774	14	CR956436	CR956436 Sus scrof	C 312	24	0.6	2139	9	F333472S14	AF333472S14 Mus muscu
C 240	25	0.6	208992	14	CR956415	CR956415 Sus scrof	C 313	24	0.6	2191	9	BC004000	BC004000 Mus muscu
C 241	25	0.6	210464	14	AC162828	AC162828 Bos tauru	C 314	24	0.6	2586	8	BC041958	BC041958 Homo sapi
C 242	25	0.6	211025	9	AC126796	AC126796 Mus muscu	C 315	24	0.6	2813	8	HSM804479	AL831368 Homo sapi
C 243	25	0.6	211380	14	CR956360	CR956360 Sus scrof	C 316	24	0.6	3053	5	AB045212	AB045212 Oreochrom
C 244	25	0.6	212859	9	AL672180	AL672180 Mouse DNA	C 317	24	0.6	4001	5	BC084325	BC084325 Xenopus l
C 245	25	0.6	212867	14	AL52071	AL52071 Bos tauru	C 318	24	0.6	4936	15	AK121600	AK121600 Oryza sat
C 246	25	0.6	212885	9	RN139121	CR937048 Rattus no	C 319	24	0.6	9417	6	AB0967	AB0967 Sequence 19
C 247	25	0.6	213808	5	AL935117	AL935117 Zebrafish	C 320	24	0.6	9417	6	A95346	A95346 Sequence 19
C 248	25	0.6	214261	14	AL125180	AL125180 Mus muscu	C 321	24	0.6	9417	6	AB03346	AB03346 Homo sapi
C 249	25	0.6	215397	8	AL355497	AL355497 Human DNA	C 322	24	0.6	19796	8	AB035346	AB035346 Homo sapi
C 250	25	0.6	220832	14	AL164224	AL164224 Bos tauru	C 323	24	0.6	31211	14	AC017253	AC017253 Drosophi1
C 251	25	0.6	222859	14	AL129158	AL129158 Rattus no	C 324	24	0.6	34575	4	AB079894	AB079894 Sus scrof
C 252	25	0.6	223477	14	CR854966	CR854966 Danio rer	C 325	24	0.6	34754	4	SSC494748	AJ94748 Sus scrof
C 253	25	0.6	225310	14	AC150170	AC150170 Gallus ga	C 326	24	0.6	40014	4	AC091538	AC091538 Sus scrof
C 254	25	0.6	226763	5	BX649352	BX649352 Zebrafish	C 327	24	0.6	40704	4	L48038	L48038 Structure o
C 255	25	0.6	226857	14	CT009542	CT009542 Sus scrof	C 328	24	0.6	43952	8	HSU10564	Z92846 Human DNA s
C 256	25	0.6	229155	8	AC018755	AC018755 Homo sapi	C 329	24	0.6	49480	8	AL450269	AL450269 Human DNA
C 257	25	0.6	230400	5	BX248506	BX248506 Zebrafish	C 330	24	0.6	62964	14	AC104959	AC104959 Homo sapi
C 258	25	0.6	238494	14	AC095717	AC095717 Rattus no	C 331	24	0.6	65389	8	AY497017	AY497017 Pongo pyg
C 259	25	0.6	239439	9	AC117573	AC117573 Mus muscu	C 332	24	0.6	65878	14	AC084777	AC084777 Homo sapi
C 260	25	0.6	240288	14	AC132298	AC132298 Mus muscu	C 333	24	0.6	66069	8	AY497016	AY497016 Pan trogl
C 261	25	0.6	242670	14	AC151187	AC151187 Bos tauru	C 334	24	0.6	66089	8	AY497014	AY497014 Gorilla g
C 262	25	0.6	245394	14	AC125756	AC125756 Rattus no	C 335	24	0.6	66091	14	AC069376	AC069376 Homo sapi
C 263	25	0.6	246542	14	AC125601	AC125601 Rattus no	C 336	24	0.6	66319	14	AC136771	AC136771 Homo sapi
C 264	25	0.6	247765	14	BX927092	BX927092 Danio rer	C 337	24	0.6	66380	14	AC130321	AC130321 Homo sapi
C 265	25	0.6	248717	14	BX470107	BX470107 Danio rer	C 338	24	0.6	68318	8	AC151408	AC151408 Pan trogl
C 266	25	0.6	251199	6	AX658289	AX658289 Sequence	C 339	24	0.6	70244	14	AC090836	AC090836 Homo sapi
C 267	25	0.6	251199	8	HUA000659	AE000659 Homo sapi	C 340	24	0.6	72456	5	BX248387	BX248387 Zebrafish
C 268	25	0.6	254530	14	AL157170	AL157170 Bos tauru	C 341	24	0.6	72690	14	AC140527_3	Continuation (4 of
C 269	25	0.6	264033	14	AC160449	AC160449 Bos tauru	C 342	24	0.6	73220	8	AY666398	AY666398 Homo sapi
C 270	25	0.6	271025	14	AC111564	AC111564 Rattus no	C 343	24	0.6	74719	14	AC139688	AC139688 Homo sapi
C 271	25	0.6	271975	14	AC111856	AC111856 Rattus no	C 344	24	0.6	75199	5	CR545472	CR545472 Zebrafish
C 272	25	0.6	275067	14	CR792438	CR792438 Danio rer	C 345	24	0.6	76815	14	AC166660	AC166660 Bos tauru
C 273	25	0.6	285748	14	AC106055	AC106055 Rattus no	C 346	24	0.6	78546	8	AY666393	AY666393 Homo sapi
C 274	25	0.6	343050	2	PF929353	AL929353 Plasmodiu	C 347	24	0.6	80601	14	AC134755	AC134755 Rattus no
C 275	24	0.6	24	6	CQ785841	CQ785841 Sequence	C 348	24	0.6	80743	8	AL162402	AL162402 Human DNA
C 276	24	0.6	24	6	CQ785843	CQ785843 Sequence	C 349	24	0.6	81466	14	AC139852	AC139852 Medicago
C 277	24	0.6	24	6	AR482289	AR482289 Sequence	C 350	24	0.6	86057	8	AC092112	AC092112 Homo sapi
C 278	24	0.6	24	6	AR482290	AR482290 Sequence	C 351	24	0.6	87738	8	BX296568	BX296568 Human DNA
C 279	24	0.6	24	6	AR482291	AR482291 Sequence	C 352	24	0.6	93240	14	AC007586	AC007586 Drosophil
C 280	24	0.6	24	6	AR482293	AR482293 Sequence	C 353	24	0.6	95642	8	AL137064	AL137064 Human DNA
C 281	24	0.6	24	6	AR482295	AR482295 Sequence	C 354	24	0.6	101064	8	CR753835	CR753835 Human DNA
C 282	24	0.6	24	6	AR482303	AR482303 Sequence	C 355	24	0.6	101385	5	CNS098BW	BX629356 Tetraodon
C 283	24	0.6	24	6	AR482305	AR482305 Sequence	C 356	24	0.6	103940	14	AP007687	AP007687 Lotus cor
C 284	24	0.6	24	6	AR482307	AR482307 Sequence	C 357	24	0.6	107436	15	AC060668	AC060668 Arabidops
C 285	24	0.6	24	6	AR482309	AR482309 Sequence	C 358	24	0.6	107898	8	CR753309	CR753309 Human DNA
C 286	24	0.6	24	6	AR482317	AR482317 Sequence	C 359	24	0.6	108736	8	CNS01RGR	AL159179 Human chr
C 287	24	0.6	24	6	AR482319	AR482319 Sequence	C 360	24	0.6	109757	9	AL162379	AL162379 Human DNA
C 288	24	0.6	199	6	CQ466406	CQ466406 Sequence	C 361	24	0.6	109929	9	AC093922	AC093922 Genomic s
C 289	24	0.6	199	6	CQ468002	CQ468002 Sequence	C 362	24	0.6	110000	1	AE017243_3	Continuation (4 of
C 290	24	0.6	341	10	AB132097	AB132097 Homo sapi	C 363	24	0.6	110000	1	AE017332_3	Continuation (5 of
C 291	24	0.6	374	6	CQ522431	CQ522431 Sequence	C 364	24	0.6	110000	1	AE017332_4	Continuation (2 of
C 292	24	0.6	467	5	AF501663	AF501663 Epinephel	C 365	24	0.6	110000	14	AC130665_1	Continuation (2 of
C 293	24	0.6	513	9	BC048542	BC048542 Mus muscu	C 366	24	0.6	110000	14	AC151717_1	Continuation (51 o
C 294	24	0.6	559	10	BV559670	BV559670 gsb23b11	C 367	24	0.6	110000	15	AP008210_050	Continuation (52 o
C 295	24	0.6	577	10	BV329768	BV329768 S241P687F	C 368	24	0.6	110000	15	AP008210_051	Continuation (52 o
C 296	24	0.6	640	10	BV288237	BV288237 S232P6229	C 369	24	0.6	110625	8	CR753301	CR753301 Human DNA
C 297	24	0.6	719	8	BC015849	BC015849 Homo sapi	C 370	24	0.6	111035	8	AC022405	AC022405 Homo sapi
C 298	24	0.6	793	6	BD107994	BD107994 EST and e	C 371	24	0.6	111437	8	AY6663401	AY6663401 Pan trogl
C 299	24	0.6	793	6	AR412441	AR412441 Sequence	C 372	24	0.6	111862	14	HSAC001237	AC001237 Homo sapi
C 300	24	0.6	793	6	AX969275	AX969275 Sequence	C 373	24	0.6	113075	8	AL512363	AL512363 Human DNA
C 301	24	0.6	804	10	BV564981	BV564981 rak56a09	C 374	24	0.6	117916	4	AC095024	AC095024 Sus scrof
C 302	24	0.6	857	10	BV575296	BV575296 G591P6076	C 375	24	0.6	118869	14	AC140573	AC140573 Macaca mu
C 303	24	0.6	891	6	BD101491	BD101491 Novel gen	C 376	24	0.6	120070	8	AC092929	AC092929 Homo sapi
C 304	24	0.6	891	6	BD021553	BD021553 Novel gen	C 377	24	0.6	120938	15	AC007017	AC007017 Arabidops
C 305	24	0.6	1434	8	BC062675	BC062675 Homo sapi	C 378	24	0.6	121483	4	BX323833	BX323833 Pig DNA s
C 306	24	0.6	1438	5	BC092772	BC092772 Danio rer	C 379	24	0.6	121909	8	HSJ492P14	HSJ492P14 Human DNA
C 307	24	0.6	1533	8	BC047881	BC047881 Homo sapi	C 380	24	0.6	122742	14	HS167P19	Z93014 Homo sapien
C 308	24	0.6	1852	8	AK024764	AK024764 Homo sapi	C 381	24	0.6	122921	8	AP005790	AP005790 Homo sapi
C 309	24	0.6	1977	6	CS033673	CS033673 Sequence	C 382	24	0.6	123414	14	AC164019	AC164019 Bos tauru
C 310	24	0.6	1977	6	CS042625	CS042625 Sequence	C 383	24	0.6	124787	8	BX248406	BX248406 Human DNA

C 384	24	0.6 126312	8	AC000026	AC000026 Homo sapi	C 457	24	0.6 166458	14	AC084737	AC084737 Homo sapi
385	24	0.6 126784	14	AV263454	AV263454 Sus scrof	458	24	0.6 166702	8	AC012594	AC012594 Homo sapi
386	24	0.6 131804	14	AC161722	AC161722 Loxodonta	C 459	24	0.6 167053	14	AC157922	AC157922 Mus muscu
C 387	24	0.6 132202	8	AC025218	AC025218 Homo sapi	C 460	24	0.6 167215	5	AY016023	AY016023 Spherooid
C 388	24	0.6 132636	14	CT009611	CT009611 Homo sapi	461	24	0.6 167358	2	AC007556	AC007556 Homo sapi
C 389	24	0.6 132875	8	AC112907	AC112907 Homo sapi	C 462	24	0.6 167368	2	AC009915	AC009915 Drosophi
C 390	24	0.6 133279	8	AC123788	AC123788 Homo sapi	C 463	24	0.6 167417	8	AC011505	AC011505 Homo sapi
C 391	24	0.6 133391	8	HSJ315G1	AL121601 Human DNA	C 464	24	0.6 167611	14	CR627494	CR627494 Danio rer
C 392	24	0.6 133786	14	AC159804	AC159804 Glycine m	C 465	24	0.6 167886	14	AC166031	AC166031 Sus scrof
C 393	24	0.6 136377	8	HSJ078N10	AL121959 Human DNA	C 466	24	0.6 168012	4	CR956391	CR956391 Pig DNA s
C 394	24	0.6 137898	5	BX571700	BX571700 Zebrafish	467	24	0.6 168506	14	AC152885	AC152885 Glycine m
C 395	24	0.6 138021	8	AC090820	AC090820 Homo sapi	468	24	0.6 168712	9	AC126929	AC126929 Mus muscu
C 396	24	0.6 138146	9	AC162298	AC162298 Mus muscu	469	24	0.6 168845	14	AC146181	AC146181 Pan trogl
C 397	24	0.6 140279	4	CR956369	CR956369 Sus scrof	470	24	0.6 169461	14	AC161828	AC161828 Mus muscu
C 398	24	0.6 141597	14	CR956416	CR956416 Pig DNA s	C 471	24	0.6 169581	2	AC008043	AC008043 Drosophi
C 399	24	0.6 144144	15	OSJ002023	AL731589 Oryza sat	C 472	24	0.6 169687	14	AC146251	AC146251 Pan trogl
C 400	24	0.6 144235	14	AC115534	AC115534 Sus scrof	473	24	0.6 170156	14	CT009670	CT009670 Sus scrof
C 401	24	0.6 145577	14	AC145762	AC145762 Sus scrof	C 474	24	0.6 170451	14	AC138167	AC138167 Sus scrof
C 402	24	0.6 146052	9	AC125140	AC125140 Mus muscu	475	24	0.6 170765	8	AL589745	AL589745 Human DNA
C 403	24	0.6 146517	14	AC069039	AC069039 Homo sapi	476	24	0.6 171165	8	AC142352	AC142352 Pan trogl
C 404	24	0.6 146543	14	AC162751	AC162751 Loxodonta	477	24	0.6 171569	8	AC132872	AC132872 Homo sapi
C 405	24	0.6 146677	8	AC087309	AC087309 Homo sapi	478	24	0.6 171623	5	AL528790	AL528790 Zebrafish
C 406	24	0.6 147663	4	SSC410870	AJ410870 Sus scrof	479	24	0.6 172252	4	BX296548	BX296548 Pig DNA s
C 407	24	0.6 149091	8	AC025917	AC025917 Homo sapi	C 480	24	0.6 172685	9	AC103941	AC103941 Mus muscu
C 408	24	0.6 149113	14	AC152859	AC152859 Sus scrof	C 481	24	0.6 172945	8	AC103725	AC103725 Homo sapi
C 409	24	0.6 149451	9	AC117191	AC117191 Mus muscu	C 482	24	0.6 173029	8	AC002059	AC002059 Homo sapi
C 410	24	0.6 149505	8	AC093767	AC093767 Homo sapi	C 483	24	0.6 173917	8	AC092810	AC092810 Homo sapi
C 411	24	0.6 149848	14	AC119516	AC119516 Rattus no	C 484	24	0.6 173982	14	AC144119	AC144119 Macaca mu
C 412	24	0.6 150638	8	AP000874	AP000874 Homo sapi	C 485	24	0.6 174187	9	AC121781	AC121781 Mus muscu
C 413	24	0.6 151277	14	RN445C15	CR937052 Rattus no	C 486	24	0.6 174700	14	AC025344	AC025344 Homo sapi
C 414	24	0.6 151383	4	CR956388	CR956388 Pig DNA s	487	24	0.6 174987	14	AC145778	AC145778 Sus scrof
C 415	24	0.6 151420	14	CR953291	CR953291 Danio rer	C 488	24	0.6 175005	14	AC137657	AC137657 Sus scrof
C 416	24	0.6 152095	9	AC154271	AC154271 Mus muscu	C 489	24	0.6 175005	14	AC137657	AC137657 Sus scrof
C 417	24	0.6 152176	8	AP003117	AP003117 Homo sapi	C 490	24	0.6 175464	8	AC069021	AC069021 Homo sapi
C 418	24	0.6 152577	14	AC148863	AC148863 Sus scrof	C 491	24	0.6 175550	14	AC009863	AC009863 Homo sapi
C 419	24	0.6 152605	14	AC136886	AC136886 Sus scrof	C 492	24	0.6 176155	14	AC012149	AC012149 Homo sapi
C 420	24	0.6 153125	14	AC148432	AC148432 Sorex ara	C 493	24	0.6 176593	8	AC092053	AC092053 Homo sapi
C 421	24	0.6 153907	8	AC146761	AC146761 Pan trogl	C 494	24	0.6 176741	8	AC016732	AC016732 Homo sapi
C 422	24	0.6 154222	14	AC061048	AC061048 Mus muscu	C 495	24	0.6 177051	14	AC138562	AC138562 Sus scrof
C 423	24	0.6 154336	8	AP006300	AP006300 Homo sapi	C 496	24	0.6 177086	5	BX470219	BX470219 Zebrafish
C 424	24	0.6 154338	9	AL669826	AL669826 Mouse DNA	C 497	24	0.6 177675	14	AC166881	AC166881 Oryctolag
C 425	24	0.6 154348	8	AC129961	AC129961 Sus scrof	C 498	24	0.6 178219	8	AC104330	AC104330 Homo sapi
C 426	24	0.6 155375	8	HS914P14	AL031117 Human DNA	C 499	24	0.6 178437	9	AC114629	AC114629 Mus muscu
C 427	24	0.6 155428	8	AC009682	AC009682 Homo sapi	C 500	24	0.6 178445	5	CR376771	CR376771 Zebrafish
C 428	24	0.6 155892	8	AL353809	AL353809 Human DNA	C 501	24	0.6 178648	14	CR956370	CR956370 Sus scrof
C 429	24	0.6 156372	5	BX072531	BX072531 Zebrafish	C 502	24	0.6 178648	14	CR956370	CR956370 Sus scrof
C 430	24	0.6 156471	14	AC138043	AC138043 Sus scrof	C 503	24	0.6 179158	14	AC140968	AC140968 Sus scrof
C 431	24	0.6 156550	14	AC015830	AC015830 Homo sapi	C 504	24	0.6 179237	14	AC026871	AC026871 Homo sapi
C 432	24	0.6 156879	14	AC026957	AC026957 Homo sapi	C 505	24	0.6 180146	14	AC150168	AC150168 Gallus ga
C 433	24	0.6 157107	14	AC140666	AC140666 Sus scrof	C 506	24	0.6 180529	9	AC154595	AC154595 Mus muscu
C 434	24	0.6 157916	9	AC110897	AC110897 Mus muscu	C 507	24	0.6 180569	14	AC145073	AC145073 Sus scrof
C 435	24	0.6 158063	4	SSC251914	AJ251914 Sus scrof	C 508	24	0.6 181137	9	AC160467	AC160467 Mus muscu
C 436	24	0.6 158837	9	AC132361	AC132361 Mus muscu	C 509	24	0.6 181244	14	AC146968	AC146968 Sus scrof
C 437	24	0.6 158971	8	AP005857	AP005857 Homo sapi	C 510	24	0.6 181336	5	BX088591	BX088591 Zebrafish
C 438	24	0.6 159322	14	AC026085	AC026085 Homo sapi	C 511	24	0.6 181605	14	AC021420	AC021420 Homo sapi
C 439	24	0.6 159442	5	AL954768	AL954768 Zebrafish	C 512	24	0.6 181605	14	AC107068	AC107068 Homo sapi
C 440	24	0.6 159671	14	AC151306	AC151306 Xenopus c	C 513	24	0.6 181922	14	AC151395	AC151395 Sus scrof
C 441	24	0.6 160275	8	AC013701	AC013701 Homo sapi	C 514	24	0.6 181936	14	AC122156	AC122156 Sus scrof
C 442	24	0.6 161141	8	AC092903	AC092903 Homo sapi	C 515	24	0.6 183163	14	AC160777	AC160777 Bos tauru
C 443	24	0.6 161551	8	AC073174	AC073174 Homo sapi	C 516	24	0.6 183521	14	AC098815	AC098815 Sus scrof
C 444	24	0.6 162112	14	AC011572	AC011572 Homo sapi	C 517	24	0.6 184157	14	AC073438	AC073438 Homo sapi
C 445	24	0.6 162335	14	AC115535	AC115535 Sus scrof	C 518	24	0.6 184995	14	CR932062	CR932062 Danio rer
C 446	24	0.6 163351	9	AL928958	AL928958 Mouse DNA	C 519	24	0.6 185081	14	CT009693	CT009693 Sus scrof
C 447	24	0.6 163355	14	AC117927	AC117927 Mus muscu	C 520	24	0.6 185954	14	AC021975	AC021975 Homo sapi
C 448	24	0.6 163943	9	AC122016	AC122016 Mus muscu	C 521	24	0.6 186090	14	AC097352	AC097352 Canis fam
C 449	24	0.6 164331	8	AC016254	AC016254 Homo sapi	C 522	24	0.6 186393	14	AC166221	AC166221 Sus scrof
C 450	24	0.6 164404	4	AC092874	AC092874 Sus scrof	C 523	24	0.6 186605	14	AC138066	AC138066 Pan trogl
C 451	24	0.6 165187	8	AP003716	AP003716 Homo sapi	C 524	24	0.6 186730	14	AC097589	AC097589 Sus scrof
C 452	24	0.6 165288	14	AC136301	AC136301 Homo sapi	C 525	24	0.6 187480	9	AC155643	AC155643 Mus muscu
C 453	24	0.6 165288	14	AC136301	AC136301 Homo sapi	C 526	24	0.6 187564	8	CNS0000E	AL049835 Human chr
C 454	24	0.6 165330	9	AC087233	AC087233 Mus Muscu	C 527	24	0.6 187800	14	CR956427	CR956427 Sus scrof
C 455	24	0.6 166002	14	AC152462	AC152462 Papio anu	C 528	24	0.6 188449	14	CR974434	CR974434 Sus scrof
C 456	24	0.6 166380	8	AC087237	AC087237 Homo sapi	C 529	24	0.6 188449	14	CR974434	CR974434 Sus scrof



C 530	24	0.6	188515	14	AC102531	AC102531 Mus muscu	603	24	0.6	231718	14	AC109552	AC109552 Rattus no
C 531	24	0.6	188725	8	AC148700	AC148700 Macaca mu	604	24	0.6	231785	14	AC094057	AC094057 Rattus no
C 532	24	0.6	189501	14	AC105436	AC105436 Sus scrof	605	24	0.6	232275	9	AC129544	AC129544 Mus muscu
C 533	24	0.6	190799	14	CT025537	CT025537 Mus muscu	606	24	0.6	233407	14	AC129452	AC129452 Rattus no
C 534	24	0.6	191726	9	AC115026	AC115026 Mus muscu	607	24	0.6	233576	14	AC113797	AC113797 Rattus no
C 535	24	0.6	191731	8	AC092139	AC092139 Homo sapi	608	24	0.6	233924	9	AC134579	AC134579 Mus muscu
C 536	24	0.6	191916	14	AC095027	AC095027 Canis fam	609	24	0.6	235407	14	AC094280	AC094280 Rattus no
C 537	24	0.6	191932	9	AC158623	AC158623 Mus muscu	610	24	0.6	235783	14	AC134202	AC134202 Rattus no
C 538	24	0.6	192752	14	AC118028	AC118028 Mus muscu	611	24	0.6	238079	14	AC096000	AC096000 Rattus no
C 539	24	0.6	192952	14	AC026966	AC026966 Homo sapi	612	24	0.6	238113	14	AC094216	AC094216 Rattus no
C 540	24	0.6	193317	8	AC098645	AC098645 Papio anu	613	24	0.6	239053	9	AC102568	AC102568 Mus muscu
C 541	24	0.6	193988	14	AC151182	AC151182 Bos tauru	614	24	0.6	240364	14	AC114875	AC114875 Rattus no
C 542	24	0.6	194130	14	AC153101	AC153101 Sus scrof	615	24	0.6	241006	14	AC152096	AC152096 Bos tauru
C 543	24	0.6	194183	8	AC026010	AC026010 Homo sapi	616	24	0.6	241531	14	AC122959	AC122959 Rattus no
C 544	24	0.6	194995	8	AC146428	AC146428 Pan trogl	617	24	0.6	241765	14	AC132557	AC132557 Rattus no
C 545	24	0.6	195159	14	AC102500	AC102500 Mus muscu	618	24	0.6	244538	14	AC105696	AC105696 Rattus no
C 546	24	0.6	195323	14	AC026612	AC026612 Homo sapi	619	24	0.6	245004	14	AC119106	AC119106 Rattus no
C 547	24	0.6	195413	9	AC068912	AC068912 Mus muscu	620	24	0.6	245556	14	AC157232	AC157232 Bos tauru
C 548	24	0.6	196798	14	AC1522196	AC1522196 Bos tauru	621	24	0.6	246654	14	AC151223	AC151223 Bos tauru
C 549	24	0.6	196859	8	AC087277	AC087277 Homo sapi	622	24	0.6	246746	14	AC121627	AC121627 Rattus no
C 550	24	0.6	197019	8	AC104435	AC104435 Homo sapi	623	24	0.6	247089	9	AC073947	AC073947 Mus muscu
C 551	24	0.6	197210	14	AC012270	AC012270 Homo sapi	624	24	0.6	247271	14	AC163918	AC163918 Bos tauru
C 552	24	0.6	197310	9	BX005240	BX005240 Mouse DNA	625	24	0.6	247786	14	AC098752	AC098752 Rattus no
C 553	24	0.6	197425	8	AC016776	AC016776 Homo sapi	626	24	0.6	248888	14	AC152225	AC152225 Bos tauru
C 554	24	0.6	197621	8	AC092809	AC092809 Homo sapi	627	24	0.6	249287	14	AC142252	AC142252 Mus muscu
C 555	24	0.6	198568	14	AC016754	AC016754 Homo sapi	628	24	0.6	250189	14	AC118814	AC118814 Rattus no
C 556	24	0.6	198573	14	AC161895	AC161895 Bos tauru	629	24	0.6	252394	2	AE014833	AE014833 Plasmodiu
C 557	24	0.6	198752	5	BX957279	BX957279 Zebrafish	630	24	0.6	252608	14	AC117146	AC117146 Rattus no
C 558	24	0.6	199694	9	AC154401	AC154401 Mus muscu	631	24	0.6	255998	14	AC125769	AC125769 Rattus no
C 559	24	0.6	199914	8	BX000009	BX000009 Pan trogl	632	24	0.6	260809	14	AC126820	AC126820 Rattus no
C 560	24	0.6	200248	14	AC145063	AC145063 Pan trogl	633	24	0.6	261832	14	AC111591	AC111591 Rattus no
C 561	24	0.6	201051	9	AC126426	AC126426 Mus muscu	634	24	0.6	265185	14	AC117637	AC117637 Mus muscu
C 562	24	0.6	201375	14	AC139789	AC139789 Homo sapi	635	24	0.6	267974	14	AC121488	AC121488 Rattus no
C 563	24	0.6	201676	14	AC105367	AC105367 Sus scrof	636	24	0.6	268795	14	AC127082	AC127082 Rattus no
C 564	24	0.6	201680	9	AC163943	AC163943 Bos tauru	637	24	0.6	270706	14	AC129862	AC129862 Rattus no
C 565	24	0.6	202338	9	AC147107	AC147107 Mus muscu	638	24	0.6	270991	14	AC136210	AC136210 Gallus ga
C 566	24	0.6	202735	14	AC159225	AC159225 Mus muscu	639	24	0.6	272855	14	AC160731	AC160731 Bos tauru
C 567	24	0.6	202998	14	AC072058	AC072058 Homo sapi	640	24	0.6	284391	14	CR974460	CR974460 Sus scrof
C 568	24	0.6	203420	9	AC159911	AC159911 Mus muscu	641	24	0.6	287397	2	AE003738	AE003738 Drosophila
C 569	24	0.6	204123	9	AC136884	AC136884 Sus scrof	642	24	0.6	294682	14	AC151287	AC151287 Mus muscu
C 570	24	0.6	204251	9	AC108430	AC108430 Mus muscu	643	24	0.6	296131	14	AL928718	AL928718 Mus muscu
C 571	24	0.6	204331	14	CR974477	CR974477 Sus scrof	644	24	0.6	328252	14	AC096800	AC096800 Rattus no
C 572	24	0.6	205285	14	CR146621	CR146621 Ootemur	645	23	0.6	23	6	CQ785842	CQ785842 Sequence
C 573	24	0.6	206042	5	BX005225	BX005225 Zebrafish	646	23	0.6	23	6	AR482299	AR482299 Sequence
C 574	24	0.6	206275	9	AC148004	AC148004 Mus muscu	647	23	0.6	23	6	AR482308	AR482308 Sequence
C 575	24	0.6	206778	14	CT009690	CT009690 Sus scrof	648	23	0.6	23	6	AX203605	AX203605 Sequence
C 576	24	0.6	207791	9	AC102501	AC102501 Mus muscu	649	23	0.6	23	6	AX613835	AX613835 Sequence
C 577	24	0.6	208067	8	AC146508	AC146508 Pan trogl	650	23	0.6	201	10	BV199973	BV199973 sqmm20284
C 578	24	0.6	208430	14	AC027641	AC027641 Homo sapi	651	23	0.6	201	10	BV208193	BV208193 sqmm22575
C 579	24	0.6	210791	8	CNS01DV7	AL133467 Human chr	652	23	0.6	292	10	BV364106	BV364106 S231P673R
C 580	24	0.6	211735	8	AC091564	AC091564 Homo sapi	653	23	0.6	301	10	AB152615	AB152615 Homo sapi
C 581	24	0.6	214578	9	AC102217	AC102217 Mus muscu	654	23	0.6	358	8	AB007148	AB007148 Homo sapi
C 582	24	0.6	214581	14	AC162042	AC162042 Bos tauru	655	23	0.6	421	6	AX401155	AX401155 Sequence
C 583	24	0.6	215342	4	AC149292	AC149292 Sus scrof	656	23	0.6	421	6	AX827297	AX827297 Sequence
C 584	24	0.6	215987	14	AC134464	AC134464 Mus muscu	657	23	0.6	458	10	HSU09A7	AL159264 STS from
C 585	24	0.6	216824	14	AC163705	AC163705 Gallus ga	658	23	0.6	460	6	CQ524582	CQ524582 Sequence
C 586	24	0.6	218379	14	AC110703	AC110703 Rattus no	659	23	0.6	481	6	BD132668	BD132668 Secretd
C 587	24	0.6	219101	14	AC152603	AC152603 Bos tauru	660	23	0.6	505	10	BV394135	BV394135 S243P6220
C 588	24	0.6	221771	14	AC162998	AC162998 Bos tauru	661	23	0.6	512	2	AF457562	AF457562 Anopheles
C 589	24	0.6	222623	14	AC097872	AC097872 Rattus no	662	23	0.6	512	10	BV215769	BV215769 S233P6418
C 590	24	0.6	223478	14	AC106224	AC106224 Rattus no	663	23	0.6	561	10	G84464	G84464 S208P6561RH
C 591	24	0.6	223552	9	AC124316	AC124316 Mus muscu	664	23	0.6	584	10	BV426869	BV426869 S237P6238
C 592	24	0.6	223728	14	AC135443	AC135443 Rattus no	665	23	0.6	604	10	BV065314	BV065314 S212P6583
C 593	24	0.6	224735	14	AC132661	AC132661 Rattus no	666	23	0.6	610	10	BV356785	BV356785 S231P6137
C 594	24	0.6	226857	14	CT009542	CT009542 Sus scrof	667	23	0.6	640	10	BV316293	BV316293 S236P6395
C 595	24	0.6	226904	14	AC111606	AC111606 Rattus no	668	23	0.6	641	10	BV023951	BV023951 S221P6643
C 596	24	0.6	227420	5	CR788254	CR788254 Zebrafish	669	23	0.6	658	10	BV222331	BV222331 S233P6276
C 597	24	0.6	228327	5	AL844514	AL844514 Zebrafish	670	23	0.6	662	4	AY285260	AY285260 Sus scrof
C 598	24	0.6	228925	14	AC162604	AC162604 Bos tauru	671	23	0.6	676	10	BV578115	BV578115 G591P6112
C 599	24	0.6	229020	14	CR974572	CR974572 Sus scrof	672	23	0.6	686	10	BV272954	BV272954 S235P682R
C 600	24	0.6	230409	9	AL772187	AL772187 Mouse DNA	673	23	0.6	688	15	AF310871	AF310871 Betula pe
C 601	24	0.6	230632	14	AC111363	AC111363 Rattus no	674	23	0.6	697	10	BV055221	BV055221 S212P604
C 602	24	0.6	230936	14	AC099285	AC099285 Rattus no	675	23	0.6	718	10	BV077076	BV077076 S212P6796

676	23	0.6	740	10	BV609693	BV609693 S215P6131	C 749	23	0.6	5565	6	AX345165
C 677	23	0.6	749	10	BV646875	BV646875 S217P6017	C 750	23	0.6	9235	9	MMU32107
C 678	23	0.6	749	10	BV102732	BV102732 MARC 2331	C 751	23	0.6	9371	2	AY600247
C 679	23	0.6	756	10	BV365552	BV365552 S231P6338	C 752	23	0.6	10480	6	AX347103
C 680	23	0.6	776	10	SSCKIIB	X56503 Porcine CKI	C 753	23	0.6	11012	4	AY044189
C 681	23	0.6	796	10	BV588904	BV588904 G591P6386	C 754	23	0.6	12486	8	AY044189
C 682	23	0.6	801	9	BC027529	BC027529 Mus muscu	C 755	23	0.6	12779	2	AE001381
C 683	23	0.6	804	9	BC038674	BC038674 Mus muscu	C 756	23	0.6	14117	6	AX331935
C 684	23	0.6	835	9	AF097438	AF097438 Mus muscu	C 757	23	0.6	14117	6	HSHA2G2EN
C 685	23	0.6	838	10	BV571202	BV571202 G591P6210	C 758	23	0.6	20486	6	AX281500
C 686	23	0.6	854	10	BV567881	BV567881 qdn91a10.	C 759	23	0.6	23369	4	BX548168
C 687	23	0.6	889	9	BC089464	BC089464 Mus muscu	C 760	23	0.6	28080	4	SCRCRYR1
C 688	23	0.6	890	9	BC078807	BC078807 Rattus no	C 761	23	0.6	28160	15	AP006092
C 689	23	0.6	925	5	BC088027	BC088027 Xenopus t	C 762	23	0.6	32196	8	AC110797
C 690	23	0.6	937	10	BV474932	BV474932 G591P6292	C 763	23	0.6	34245	6	CQ869707
C 691	23	0.6	965	9	BC027755	BC027755 Mus muscu	C 764	23	0.6	35376	5	BX890563
C 692	23	0.6	1023	9	BC024400	BC024400 Mus muscu	C 765	23	0.6	41613	8	AC004790
C 693	23	0.6	1046	2	AY058495	AY058495 Drosophill	C 766	23	0.6	42210	8	AC112724
C 694	23	0.6	1061	15	AY065275	AY065275 Arabidops	C 767	23	0.6	45154	14	AC164871
C 695	23	0.6	1205	2	AK115293	AK115293 Clona int	C 768	23	0.6	45653	5	BX323016
C 696	23	0.6	1229	9	BC058805	BC058805 Mus muscu	C 769	23	0.6	48042	14	AC020217
C 697	23	0.6	1319	8	BC032031	BC032031 Homo sapi	C 770	23	0.6	49705	8	AC087750
C 698	23	0.6	1320	9	BC100465	BC100465 Mus muscu	C 771	23	0.6	50768	8	BS000149
C 699	23	0.6	1342	5	BC094397	BC094397 Xenopus 1	C 772	23	0.6	53301	14	AC164791
C 700	23	0.6	1421	9	AF180805	AF180805 Mus muscu	C 773	23	0.6	54710	9	AL928833
C 701	23	0.6	1460	8	AB056757	AB056757 Macaca fa	C 774	23	0.6	58864	15	AP000371
C 702	23	0.6	1492	6	CQ976593	CQ976593 Sequence	C 775	23	0.6	60079	8	AL596284
C 703	23	0.6	1492	6	AF547296	AF547296 Homo sapi	C 776	23	0.6	61697	8	AL354933
C 704	23	0.6	1566	8	AF247788	AF247788 Homo sapi	C 777	23	0.6	62102	4	CR974593
C 705	23	0.6	1567	4	AY484494S9	AY484494 Sus scrofa	C 778	23	0.6	62868	14	AC016799
C 706	23	0.6	1610	8	BC007548	BC007548 Homo sapi	C 779	23	0.6	62918	14	AL160289
C 707	23	0.6	1610	8	BC033123	BC033123 Homo sapi	C 780	23	0.6	63761	14	AC117737
C 708	23	0.6	1642	8	BC015353	BC015353 Homo sapi	C 781	23	0.6	64477	14	AC117554
C 709	23	0.6	1666	15	BT002864	BT002864 Arabidops	C 782	23	0.6	67124	8	AC026415
C 710	23	0.6	1684	4	AY864618	AY864618 Sus scrofa	C 783	23	0.6	67718	14	AC100706
C 711	23	0.6	1780	9	BC062053	BC062053 Rattus no	C 784	23	0.6	69060	8	AC005196
C 712	23	0.6	1821	9	BC027238	BC027238 Mus muscu	C 785	23	0.6	69699	4	SSCS04726
C 713	23	0.6	1947	9	AB116526	AB116526 Rattus no	C 786	23	0.6	69784	5	BX537312
C 714	23	0.6	1977	8	HSM805868	BX537781 Homo sapi	C 787	23	0.6	72382	8	AC110082
C 715	23	0.6	1980	9	BC004042	BC004042 Mus muscu	C 788	23	0.6	72554	4	AL773521
C 716	23	0.6	1983	8	BC03063	BC03063 Homo sapi	C 789	23	0.6	77162	8	AC112694
C 717	23	0.6	1985	9	BC079044	BC079044 Rattus no	C 790	23	0.6	77190	8	AC112694
C 718	23	0.6	1991	8	AK128768	AK128768 Homo sapi	C 791	23	0.6	78592	14	AC026589
C 719	23	0.6	2015	8	AB051127	AB051127 Macaca fa	C 792	23	0.6	79003	9	BX119960
C 720	23	0.6	2052	9	BC078996	BC078996 Rattus no	C 793	23	0.6	80015	8	AC110085
C 721	23	0.6	2337	2	AY061276	AY061276 Drosophill	C 794	23	0.6	80636	8	HS384D21
C 722	23	0.6	2424	8	HSU06669	U60669 Human 1 alp	C 795	23	0.6	81357	8	AC005792
C 723	23	0.6	2433	9	BC022143	BC022143 Mus muscu	C 796	23	0.6	81971	8	HS593C16
C 724	23	0.6	2458	9	BC026615	BC026615 Mus muscu	C 797	23	0.6	82771	14	AC011953
C 725	23	0.6	2472	9	BC013552	BC013552 Mus muscu	C 798	23	0.6	83029	8	AC022127
C 726	23	0.6	2479	5	BC091060	BC091060 Xenopus t	C 799	23	0.6	84495	9	BX537352
C 727	23	0.6	2543	8	BC010144	BC010144 Homo sapi	C 800	23	0.6	84984	14	AC165468
C 728	23	0.6	2645	9	AB107226	AB107226 Cricetulu	C 801	23	0.6	85712	8	AC141556
C 729	23	0.6	2654	9	BC051254	BC051254 Mus muscu	C 802	23	0.6	90877	8	AC069436
C 730	23	0.6	2681	8	AB178966	AB178966 Macaca fa	C 803	23	0.6	91912	8	AL589935
C 731	23	0.6	2723	9	MMU401376	MMU401376 Mus muscu	C 804	23	0.6	92245	8	AC073968
C 732	23	0.6	2778	13	LPDR13	LPDR13 La France d	C 805	23	0.6	93578	8	HSAC00055
C 733	23	0.6	2790	8	BC038220	BC038220 Homo sapi	C 806	23	0.6	93800	8	HSJ837021
C 734	23	0.6	2798	2	BT011496	BT011496 Drosophill	C 807	23	0.6	94387	8	AL136532
C 735	23	0.6	2843	9	BC062011	BC062011 Rattus no	C 808	23	0.6	94647	15	AC145372
C 736	23	0.6	3145	5	AJ720266	AJ720266 Gallus ga	C 809	23	0.6	94656	15	AC138579
C 737	23	0.6	3197	9	AF041431	AF041431 Mus muscu	C 810	23	0.6	94893	8	AC064860
C 738	23	0.6	3382	9	RATCARCPT	LA4617 Rattus norv	C 811	23	0.6	95356	8	AL139112
C 739	23	0.6	3456	15	AK120832	AK120832 Oryza sat	C 812	23	0.6	95565	8	AC000379
C 740	23	0.6	3501	15	AK121669	AK121669 Oryza sat	C 813	23	0.6	96206	8	AC013283
C 741	23	0.6	3510	5	AJ719395	AJ719395 Gallus ga	C 814	23	0.6	96792	8	HSU95741
C 742	23	0.6	3741	5	AB185847	AB185847 Oryzias 1	C 815	23	0.6	97325	9	AL928836
C 743	23	0.6	3802	8	BC096750	BC096750 Homo sapi	C 816	23	0.6	98765	14	AC165807
C 744	23	0.6	3894	9	BC053925	BC053925 Mus muscu	C 817	23	0.6	101224	8	AC106750
C 745	23	0.6	4117	9	BC063078	BC063078 Mus muscu	C 818	23	0.6	102298	9	AL671671
C 746	23	0.6	4399	15	AK120928	AK120928 Oryza sat	C 819	23	0.6	103004	8	HS171H20
C 747	23	0.6	5107	8	HS283334	Z83334 H.sapiens R	C 820	23	0.6	103510	14	AP007944
C 748	23	0.6	5529	8	HSR833AGE	X87373 Homo sapien	C 821	23	0.6	104492	14	AC124910

822	23	0.6 104739	14	AC161817_8	Continuation (9 of	c 895	23	0.6 137520	4	EA459965	AJ459965 Equus cab
823	23	0.6 105171	8	AY663410	AY663410 Homo sapi	c 896	23	0.6 137737	8	AC010378	AC010378 Homo sapi
824	23	0.6 106857	15	AC157777	AC157777 Medicago	c 897	23	0.6 137783	8	HS564M11	AC035409 Human DNA
825	23	0.6 107044	14	AC142570	AC142570 Gallus ga	c 898	23	0.6 139771	15	AC135564	AC135564 Oryza sat
826	23	0.6 107132	14	AF312914	AF312914 Homo sapi	c 899	23	0.6 140112	14	AC148127	AC148127 Otolomur
827	23	0.6 107593	14	AC107484	AC107484 Homo sapi	c 900	23	0.6 140211	8	HS0J12208	AL096678 Human DNA
828	23	0.6 107910	6	CS086315	CS086315 Sequence	c 901	23	0.6 140773	14	AC012227	AL096678 Human DNA
829	23	0.6 107910	8	HUAC00394	CS086315 Sequence	c 902	23	0.6 141200	8	AC012227	AL096678 Human DNA
830	23	0.6 109441	8	AY663404	AY663404 Homo sapi	c 903	23	0.6 141200	8	AC012227	AL096678 Human DNA
831	23	0.6 110000	1	BA0017244_06	Continuation (7 of	c 904	23	0.6 141293	14	AP004337	AP004337 Oryza sat
832	23	0.6 110000	1	BA0017244_06	Continuation (7 of	c 904	23	0.6 141380	14	AC163972	AC163972 Loxodonta
833	23	0.6 110000	30	BA000037_30	Continuation (31 of	c 905	23	0.6 141620	14	AC142444	AC142444 Homo sapi
834	23	0.6 110000	8	AB128049_03	Continuation (4 of	c 906	23	0.6 142446	15	AP005126	AP005126 Oryza sat
835	23	0.6 110000	8	AB128049_04	Continuation (5 of	c 907	23	0.6 142730	14	AC108186	AC108186 Felis cat
836	23	0.6 110000	8	AB128049_05	Continuation (6 of	c 908	23	0.6 142759	14	AC163267	AC163267 Rhinoloph
837	23	0.6 110000	14	AC109710_3	Continuation (4 of	c 909	23	0.6 142790	14	AC009616	AC009616 Homo sapi
838	23	0.6 110000	14	AC140148_1	Continuation (2 of	c 910	23	0.6 142903	9	AC117253	AC117253 Mus muscu
839	23	0.6 110000	14	AC145312_2	Continuation (3 of	c 911	23	0.6 142971	8	AC020629	AC020629 Homo sapi
840	23	0.6 110000	14	BX119990_1	Continuation (2 of	c 912	23	0.6 143644	14	AC149903	AC149903 Strongylo
841	23	0.6 110000	14	CR974431_0	CR974431 Sub scrof	c 913	23	0.6 143750	8	AC109999	AC109999 Homo sapi
842	23	0.6 110000	14	CR974431_3	Continuation (4 of	c 914	23	0.6 143947	8	AC012462	AC012462 Homo sapi
843	23	0.6 110000	15	AP008217_007	Continuation (8 of	c 915	23	0.6 144026	14	AC165403	AC165403 Callicebu
844	23	0.6 110000	15	AP008217_218	Continuation (219	c 916	23	0.6 144034	14	AC069244	AC069244 Homo sapi
845	23	0.6 110000	15	AP008207_062	Continuation (63 o	c 917	23	0.6 144446	8	AC146219	AC146219 Pan trogl
846	23	0.6 110000	15	AP008209_217	Continuation (218	c 918	23	0.6 144500	14	AC010231	AC010231 Homo sapi
847	23	0.6 110000	15	AP008213_219	Continuation (220	c 919	23	0.6 145417	14	AC138169	AC138169 Sub scrof
848	23	0.6 110541	8	AC107061	AC107061 Homo sapi	c 920	23	0.6 145633	8	BS000107	BS000107 Pan trogl
849	23	0.6 112259	14	AC026579	AC026579 Homo sapi	c 921	23	0.6 146455	8	AC127512	AC127512 Homo sapi
850	23	0.6 112568	8	AC073886	AC073886 Homo sapi	c 922	23	0.6 147021	8	AC006501	AC006501 Homo sapi
851	23	0.6 113114	8	AC079400	AC079400 Homo sapi	c 923	23	0.6 147288	8	AL355812	AL355812 Human DNA
852	23	0.6 114156	9	AL590390	AL590390 Mouse DNA	c 924	23	0.6 147601	14	AC012655	AC012655 Homo sapi
853	23	0.6 114180	5	CR936359	CR936359 Zebrafish	c 925	23	0.6 147854	8	AC021422	AC021422 Homo sapi
854	23	0.6 116609	5	BS111244	BS111244 Zebrafish	c 926	23	0.6 147862	8	AC108922	AC108922 Homo sapi
855	23	0.6 117872	8	AL162396	AL162396 Human DNA	c 927	23	0.6 147978	8	AL137781	AL137781 Human DNA
856	23	0.6 118276	8	AC004148	AC004148 Homo sapi	c 928	23	0.6 148468	14	AL355882	AL355882 Homo sapi
857	23	0.6 118404	4	AL773591	AL773591 Pig DNA s	c 929	23	0.6 148620	9	BX890605	BX890605 Mouse DNA
858	23	0.6 120180	5	BS233592	BS233592 Zebrafish	c 930	23	0.6 148870	8	HS173A13	AL355688 Human DNA
859	23	0.6 120989	14	AC156746	AC156746 Bos tauru	c 931	23	0.6 149180	8	AC114781	AC114781 Homo sapi
860	23	0.6 121223	15	AC123514	AC123514 Oryza sat	c 932	23	0.6 149188	8	AC114781	AC114781 Homo sapi
861	23	0.6 121780	8	HS774124	AL031290 Human DNA	c 933	23	0.6 149494	14	AC161411	AC161411 Mus muscu
862	23	0.6 121936	9	AL691513	AL691513 Mouse DNA	c 934	23	0.6 149562	14	AC143353	AC143353 Homo sapi
863	23	0.6 122167	5	BS111009	BS111009 Zebrafish	c 935	23	0.6 149859	14	AC148356	AC148356 Sorex ara
864	23	0.6 122825	8	AC108067	AC108067 Homo sapi	c 936	23	0.6 149902	14	AC110667	AC110667 Canis fam
865	23	0.6 123782	8	HS0J756N5	AL132825 Human DNA	c 937	23	0.6 150005	8	AL450344	AL450344 Human DNA
866	23	0.6 125042	14	AC164944	AC164944 Sub scrof	c 938	23	0.6 150022	8	AC122536	AC122536 Mus muscu
867	23	0.6 126867	8	AC022443	AC022443 Homo sapi	c 939	23	0.6 150122	9	BX088567	BX088567 Mouse DNA
868	23	0.6 126933	14	AC148351	AC148351 Sorex ara	c 940	23	0.6 150128	14	AC166227	AC166227 Sorex ara
869	23	0.6 127234	4	AB206998	AB206998 Sub scrof	c 941	23	0.6 150159	4	AB192510	AB192510 Sub scrof
870	23	0.6 128073	9	AL772151	AL772151 Mouse DNA	c 942	23	0.6 150409	9	AC122911	AC122911 Mus muscu
871	23	0.6 128150	14	AC148748	AC148748 Sorex ara	c 943	23	0.6 150649	9	AC102918	AC102918 Mus muscu
872	23	0.6 128243	14	AC162129	AC162129 Loxodonta	c 944	23	0.6 151194	9	AC112988	AC112988 Mus muscu
873	23	0.6 128990	8	AC113344	AC113344 Homo sapi	c 945	23	0.6 151383	4	CR956388	CR956388 Pig DNA s
874	23	0.6 129264	4	CR956409	CR956409 Pig DNA s	c 946	23	0.6 151405	15	AP004261	AP004261 Oryza sat
875	23	0.6 129624	14	BX296540	BX296540 Sub scrof	c 947	23	0.6 151523	4	AC147901	AC147901 Xenopus t
876	23	0.6 130117	9	RN394C12	CR937051 Rattus no	c 948	23	0.6 152058	4	AC091756	AC091756 Sub scrof
877	23	0.6 130267	14	AC010597	AL031012 Human DNA	c 949	23	0.6 152211	9	SSC251829	AC251829 Sub scrof
878	23	0.6 130988	8	HS364I22	AL031012 Human DNA	c 950	23	0.6 152279	9	AC155915	AC155915 Mus muscu
879	23	0.6 131005	14	AC150890	AC150890 Medicago	c 951	23	0.6 153438	14	AC117959	AC117959 Rattus no
880	23	0.6 131621	14	AC166224	AC166224 Sorex ara	c 952	23	0.6 153745	4	AY587061	AY587061 Sub scrof
881	23	0.6 131860	9	AC133163	AC133163 Mus muscu	c 953	23	0.6 153745	4	AY587061	AY587061 Sub scrof
882	23	0.6 132150	8	AC005586	AC005586 Homo sapi	c 954	23	0.6 153758	14	AC022451	AC022451 Homo sapi
883	23	0.6 132421	9	AC116569	AC116569 Mus muscu	c 955	23	0.6 153857	14	AF336380	AF336380 Mus muscu
884	23	0.6 133538	14	AC160877	AC160877 Aterlix	c 956	23	0.6 153866	8	AC009307	AC009307 Homo sapi
885	23	0.6 133836	8	AC101343	AC101343 Homo sapi	c 957	23	0.6 154101	14	AC136885	AC136885 Sub scrof
886	23	0.6 133838	8	AC105410	AC105410 Homo sapi	c 958	23	0.6 154513	14	AC145047	AC145047 Sub scrof
887	23	0.6 134926	8	AP001347	AP001347 Homo sapi	c 959	23	0.6 154755	14	CR974570	CR974570 Sub scrof
888	23	0.6 135412	14	AC148279	AC148279 Sorex ara	c 960	23	0.6 154954	8	AC022507	AC022507 Homo sapi
889	23	0.6 135508	14	AC120215	AC120215 Sub scrof	c 961	23	0.6 155263	14	BS927284	BS927284 Danio rer
890	23	0.6 135756	14	AC158274	AC158274 Aterlix	c 962	23	0.6 155332	14	AC040900	AC040900 Homo sapi
891	23	0.6 136646	14	AC135220	AC135220 Sub scrof	c 963	23	0.6 155362	14	AC148434	AC148434 Sorex ara
892	23	0.6 137001	8	AP001005	AP001005 Homo sapi	c 964	23	0.6 155367	9	AC154778	AC154778 Mus muscu
893	23	0.6 137005	14	AC104678	AC104678 Sub scrof	c 965	23	0.6 155512	14	AC148909	AC148909 Otolomur
894	23	0.6 137371	9	AC115124	AC115124 Mus muscu	c 966	23	0.6 155663	14	AC052857	AC052857 Aterlix
	23					c 967	23	0.6 155710	4	AC092875	AC092875 Sub scrof

c 968	23	0.6	155992	14	AC083854	AC083854 Homo sapi
c 969	23	0.6	155992	14	AC034184	AC034184 Homo sapi
c 970	23	0.6	156328	14	AC148280	AC148280 Orex ara
c 971	23	0.6	156331	15	AC144700	AC144700 Oryza sat
c 972	23	0.6	156678	14	AC145026	AC145026 Sus scrof
c 973	23	0.6	156700	14	AC027051	AC027051 Homo sapi
c 974	23	0.6	156818	8	AC104463	AC104463 Homo sapi
c 975	23	0.6	156945	9	AC105944	AC105944 Mus muscu
c 976	23	0.6	157248	8	AC022312	AC022312 Human Chr
c 977	23	0.6	157654	14	AC091791	AC091791 Sus scrof
c 978	23	0.6	157870	4	AB192509	AB192509 Sus scrof
c 979	23	0.6	157965	14	AC011220	AC011220 Homo sapi
c 980	23	0.6	158423	14	AC114299	AC114299 Homo sapi
c 981	23	0.6	158527	14	AC027640	AC027640 Homo sapi
c 982	23	0.6	158608	8	AC051635	AC051635 Homo sapi
c 983	23	0.6	159150	14	AC022863	AC022863 Homo sapi
c 984	23	0.6	159350	14	AC136882	AC136882 Sus scrof
c 985	23	0.6	160429	5	BX004784	BX004784 Zebrafish
c 986	23	0.6	160460	14	AC144692	AC144692 Sus scrof
c 987	23	0.6	160574	9	AC140192	AC140192 Mus muscu
c 988	23	0.6	160831	9	AL732450	AL732450 Mouse DNA
c 989	23	0.6	160951	14	AC112181	AC112181 Homo sapi
c 990	23	0.6	161180	14	AC146771	AC146771 Canis fam
c 991	23	0.6	161309	8	AC100839	AC100839 Homo sapi
c 992	23	0.6	161823	4	AC097231	AC097231 Sus scrof
c 993	23	0.6	162029	14	AC145046	AC145046 Sus scrof
c 994	23	0.6	162029	14	AC145046	AC145046 Sus scrof
c 995	23	0.6	162065	14	AC117061	AC117061 Rattus no
c 996	23	0.6	162111	14	AC138170	AC138170 Sus scrof
c 997	23	0.6	162111	14	AC138170	AC138170 Sus scrof
c 998	23	0.6	162405	8	AC002352	AC002352 Homo sapi
c 999	23	0.6	162681	8	AC087835	AC087835 Homo sapi
c1000	23	0.6	162770	9	AC121904	AC121904 Mus muscu

## ALIGNMENTS

RESULT 1  
 AR482278 3919 bp DNA linear PAT 14-MAY-2004  
 LOCUS Sequence 1 from patent US 6703197.  
 DEFINITION AR482278  
 ACCESSION AR482278.1 GI:47244567  
 VERSION  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 1 (bases 1 to 3919)  
 Gravel,R.A., Rozen,R., Leclerc,D., Goyette,P. and Campeau,E.  
 Human methionine synthase: cloning, and methods for evaluating risk  
 of neural tube defects, cardiovascular disease, and cancer  
 Patent: US 6703197-A 1 09-MAR-2004;  
 Martine R&D, Inc.; Montreal;  
 CAX;  
 FEATURES  
 Location/Qualifiers  
 1..3919  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 3919; DB 6; Length 3919;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGTACCTGTGGAGACGCTTCTCTGCGCGCCCTCTGCGAAGGAGGAGACTCGAC 60  
 Db 1 GGTACCTGTGGAGACGCTTCTCTGCGCGCCCTCTGCGAAGGAGGAGACTCGAC 60  
 Qy 61 AACATGTCAACCGCGCTCCAAGACTGTGCGCAACCGCAAGCTCTGAAGAAACCTCGG 120  
 Db 61 AACATGTCAACCGCGCTCCAAGACTGTGCGCAACCGCAAGCTCTGAAGAAACCTCGG 120

121 GATGAGATCAATGCGCATTCCTGCGAAGAGGATATATGCTGCTGGATGGAGGATGGGACC 180  
 121 GATGAGATCAATGCGCATTCCTGCGAAGAGGATATATGCTGCTGGATGGAGGATGGGACC 180  
 181 ATGATCCAGCGGAGAGAGCTTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCAT 240  
 181 ATGATCCAGCGGAGAGAGCTTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCAT 240  
 241 GCGAGGCGCTCAAGAGGCAACATGACATTTTAAAGTATTAACCTCAGCTCATGTCATTAC 300  
 241 GCGAGGCGCTCAAGAGGCAACATGACATTTTAAAGTATTAACCTCAGCTCATGTCATTAC 300  
 301 CAAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGC 360  
 301 CAAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGC 360  
 361 AGCATTAGTATGCGCAAGCTGATATGAGCTTGAACAATGCGGCTTACCGGATGAACATG 420  
 361 AGCATTAGTATGCGCAAGCTGATATGAGCTTGAACAATGCGGCTTACCGGATGAACATG 420  
 421 TGGCTGCGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGACACGGAATTAAG 480  
 421 TGGCTGCGAGGAGTGGCCAGAAAGCTGCGGAGGAGTAACTCTCCAGACACGGAATTAAG 480  
 481 AGGTTTGTGCGAGGAGCTCTGCGTCCGACTAATAGACACTCTCTGTGTCCCATCTGTG 540  
 481 AGGTTTGTGCGAGGAGCTCTGCGTCCGACTAATAGACACTCTCTGTGTCCCATCTGTG 540  
 541 GAAAGCCCGGATTTATAGAACATCACTTTGATAGCTTTGTTGAAGCATACCAAGAGCAG 600  
 541 GAAAGCCCGGATTTATAGAACATCACTTTGATAGCTTTGTTGAAGCATACCAAGAGCAG 600  
 601 GCCAAGGAGCTTCTGATGCGGGGTTGATCTTACTCATTTGAAACTATTTTGATCT 660  
 601 GCCAAGGAGCTTCTGATGCGGGGTTGATCTTACTCATTTGAAACTATTTTGATCT 660  
 661 GCCAATGCCAAGGAGCGCTTTGTTGCACTCCAAAATCTTTTGGAGAGAAATATGCTCCC 720  
 661 GCCAATGCCAAGGAGCGCTTTGTTGCACTCCAAAATCTTTTGGAGAGAAATATGCTCCC 720  
 721 CGGCTTATCTTTATTTTTCAGGAGCGATCGTTGATAAAAGTGGGCGGACTCTTTCCGAC 780  
 721 CGGCTTATCTTTATTTTTCAGGAGCGATCGTTGATAAAAGTGGGCGGACTCTTTCCGAC 780  
 781 ACAGGAGGAGGATTTCTCATCGGCTCTCTCATGGAGAACCACTCTGCAATGATTAAT 840  
 781 ACAGGAGGAGGATTTCTCATCGGCTCTCTCATGGAGAACCACTCTGCAATGATTAAT 840  
 841 TGTGCTTTGGGTCAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATGTTACACA 900  
 841 TGTGCTTTGGGTCAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATGTTACACA 900  
 901 GCCTATGCTCTCTGTTATCCAAATGAGGCTTTCCCAACACCTTTGGTGAATGATGAA 960  
 901 GCCTATGCTCTCTGTTATCCAAATGAGGCTTTCCCAACACCTTTGGTGAATGATGAA 960  
 961 ACCTCTTCTATGATGCGCAAGCACTTAAAGATTTTGTCTATGATGCTTGGTCAATATA 1020  
 961 ACCTCTTCTATGATGCGCAAGCACTTAAAGATTTTGTCTATGATGCTTGGTCAATATA 1020  
 1021 GTTGGAGGATGCTGTGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
 1021 GTTGGAGGATGCTGTGGTCAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
 1081 AATTGTAAGCTTAGAGTTCCACTGCTGCTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
 1081 AATTGTAAGCTTAGAGTTCCACTGCTGCTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
 1141 CTAGAGCCCTTCAGGATTTGACCGGTACACCACTTTTGTAACTTGGAGAGCGCTGTAAT 1200  
 1141 CTAGAGCCCTTCAGGATTTGACCGGTACACCACTTTTGTAACTTGGAGAGCGCTGTAAT 1200  
 1201 GTTSCAGGATCAAGGAAGTTTGTCTAACTCATCATGCGAGGAAACTATGAAGAAGCCTTG 1260



CDS	64..3861	/note="methionine synthase; cobalamin binding protein; cytosolic protein"	/codon_start=1	/product="5-methyltetrahydrofolate-homocysteine methyltransferase"	/protein_id="AAC51188.1"	/db_xref="GI:1923221"	/translation="MSPALQDLSQPEGLKKTLDENAILQKRWLVLDGGMGTMIQRE KLNBEHFRGQFQKHARPLKNGNDILSIQPDVIYQIHKEYVLLAGADIIETNFSSTIS IAQDYGLEHLYARMNCSAGVAKAEVTLQGTIKFVAGALQNTKLTLSVSPVE TAQDYGLERHLYARMNCSAGVAKAEVTLQGTIKFVAGALQNTKLTLSVSPVE RPDYRNITDELVEAYQEOAGLGGVDILIIETIFDTANAKAALFALQNLPEEKYA RPIFTSGTIVDKSGRTSGGTGEGFVLSVSHGPEPLCIGLNCALGAEMRPFPIEIKG CTTAYVLCYNAGLNTGCDYDTPSMMAKHLKDFAMQGLVNI VGGCCGSTDPIREI AEAVNCKRPVPATAPAFEGHMLLSGLEPPRIGPYTNFNIGERCNVAGSRKAKLMA GNYEALCVAKVQVEMGAQVLNMDMGDPSAMTRFCNLIASEPDIKAVPLCLIDS SNPTVEAGLKCCKGKCVNSISLKEGEDDFLEKARKIKKYGAAMVMAFDEGQATE TDKIRVCITRAYHLVKKLGNPNDIIFDNLITITGMBEHNLYAINFIHATKVIKE TLLGARISGGLSNLGSFRGMEAIRAMHGVLVYHAIKSGMDMGI VNAQNLVYDDIH KELLQOLCEDLIWKDPEATEKLLRYAOTGTGCKKVIOTDEWRNGVBERLYALVKG ISEKHIEDTEARLNQKPYRPLNIIIEGLPMNGMKIIVGLFCGAKMFL PQVLSARVM KXAVHLIIPFMEKEEREKRVLTNGTVEEDPYGTITVLATVKGVDHDIKNIYGVVLGC NFRVYLDGVTMPCDKILKAALDHKADIIGLSGLITPSLDEMI FVAKEMERLAIRLP LDIQGHYESLKERRYPLPSQARKSGFQMDLSEPHVKPTFTGTQVDFDYDLQKLVYD DIRQKPPFQVMOQLRGKVPNRGFPKI FNDKTVGGEARKVYDDAHNMLNTLISOKKLRRG VVGFPQAQIQDDIHLVAAEAAVPOAAEPIATPYGLRQAKDSASTEPYCYCLSDFIAP LHSGIRDYGLFVAVACFGVEELSKAYEDDGDYSSIMVKAALGDRLAEAFABELHVR RLWAYCGSEQLDVADRLRLRYKIRPAPGPSQPDHTEKLTMWLADIETQSGIRLT ESLAMAPASAVSGLYFSLNLSKYFAVGKISKQVVEDYALRKNIISVAEVEKMLGPILY DTD"
	misc_feature	777..876	/note="encodes B12 binding region"				
	variation	2703..2705	/note="deletion of isoleucine (1881) in cblG patient"				
	variation	2819	/replace="g"				
	variation	2821	/note="D919G polymorphism site; creates conversion of aspartic acid to glycine"				
	variation	2821	/replace="g"				
	variation	2821	/note="mutation H920D in cblG patient; creates conversion of histidine to aspartic acid"				
	variation	2821	/replace="g"				
	ORIGIN	Query Match 97.6%; Score 3823; DB 8; Length 3917; Best Local Similarity 99.9%; Pred. No. 0; Matches 3873; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
	Qy	2	GTCACTGTGGAGACGCTCTTCTCTCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	61			
	Db	2	GTCACTGTGGAGACGCTCTTCTCTCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	61			
	Qy	62	ACATGTGACCCCGCGCTCCAAGACCTGTGCGCAACCGAAGGCTTGAAGAAAACCCCTGCGG	121			
	Db	62	ACATGTGACCCCGCGCTCCAAGACCTGTGCGCAACCGAAGGCTTGAAGAAAACCCCTGCGG	121			
	Qy	122	ATGAGATCAATGCCATTCTTCAGAAGAGGATTATGTGCTGGATGGAGGATGGGGACCA	181			
	Db	122	ATGAGATCAATGCCATTCTTCAGAAGAGGATTATGTGCTGGATGGAGGATGGGGACCA	181			
	Qy	182	TGATCCAGCGGGAGAGCTTAAACGAGAACACTTCCGAGGCTCAGGAATTTAAAGATCATG	241			
	Db	182	TGATCCAGCGGGAGAGCTTAAACGAGAACACTTCCGAGGCTCAGGAATTTAAAGATCATG	241			
	Qy	242	CCAGGCGCGCTGAAAGGCAACAAATGACATTTTAAAGTATAACTCAGCCCTGATGTCTATTACC	301			
	Db	242	CCAGGCGCGCTGAAAGGCAACAAATGACATTTTAAAGTATAACTCAGCCCTGATGTCTATTACC	301			
	Qy	302	AAATCCATAAGGAATACTTCTGCTGGCGGAGATATCATATGAAACAAATATCTTTTAGCA	361			
	Db	302	AAATCCATAAGGAATACTTCTGCTGGCGGAGATATCATATGAAACAAATATCTTTTAGCA	361			



QY 362 GCACATGATATGCCCCAAGCTGACTATGAGCTTTGAAACACTTGGCCTACCGGATGAACATGT 421  
DB 362 GCACATGATATGCCCCAAGCTGACTATGAGCTTTGAAACACTTGGCCTACCGGATGAACATGT 421  
QY 422 GCTCTCAGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGAGAAATTAAGA 481  
DB 422 GCTCTCAGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGAGAAATTAAGA 481  
QY 482 GGTTCGTGCGAGGGCTCTGGGTCCGACTTAATAAGACACTCTCTGTGTCCCACTCTGTGG 541  
DB 482 GGTTCGTGCGAGGGCTCTGGGTCCGACTTAATAAGACACTCTCTGTGTCCCACTCTGTGG 541  
QY 542 AAAGCCGGAATATAGGAACATCACAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
DB 542 AAAGCCGGAATATAGGAACATCACAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
QY 602 CCAAGGACTCTTGATGAGGGGTTGATCTTACTCATTTGAACATAATTTTGTGATGACTG 661  
DB 602 CCAAGGACTCTTGATGAGGGGTTGATCTTACTCATTTGAACATAATTTTGTGATGACTG 661  
QY 662 CCAATGCCAAGGAGCCTTTGTTGCACTCCAAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
DB 662 CCAATGCCAAGGAGCCTTTGTTGCACTCCAAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
QY 722 GGCCTATCTTTATTTCAAGGACGATCGTTGATAAAGTGGGGGACTCTTTCCGGACAGA 781  
DB 722 GGCCTATCTTTATTTCAAGGACGATCGTTGATAAAGTGGGGGACTCTTTCCGGACAGA 781  
QY 782 CAGGAGAGGATTTGTCATCAGGGTCTCATGAGAACCACTCTGCATTTGATGATTAATTT 841  
DB 782 CAGGAGAGGATTTGTCATCAGGGTCTCATGAGAACCACTCTGCATTTGATGATTAATTT 841  
QY 842 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATATGTAACACAG 901  
DB 842 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATATGTAACACAG 901  
QY 902 CTTATGCTCTCTGTTATCCAAATGCAAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 961  
DB 902 CTTATGCTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 961  
QY 962 CGCCTTCTATGATGGCCAGACCTTAAGAGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
DB 962 CGCCTTCTATGATGGCCAGACCTTAAGAGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
QY 1022 TTGGAGATGCTGTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
DB 1022 TTGGAGATGCTGTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
QY 1082 ATTGTAAAGCTTAGATTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141  
DB 1082 ATTGTAAAGCTTAGATTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141  
QY 1142 TAGAGCCCTTCAGGATGACCGTACACCAACTTTGTTTAACTATGAGAGCCGTGTAAATG 1201  
DB 1142 TAGAGCCCTTCAGGATGACCGTACACCAACTTTGTTTAACTATGAGAGCCGTGTAAATG 1201  
QY 1202 TTGCAGGATCAAGGAATTTGCTTAAACTCATATGCGAGAACTATGAAGAGCCCTGT 1261  
DB 1202 TTGCAGGATCAAGGAATTTGCTTAAACTCATATGCGAGAACTATGAAGAGCCCTGT 1261  
QY 1262 GTGTTGCCAAAGTGCAAGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATG 1321  
DB 1262 GTGTTGCCAAAGTGCAAGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATG 1321  
QY 1322 GCATGCTAGATGCTCAAGTGCATGACAGAGATTTTGAACCTTAATTTGCTCCGAGCCAG 1381  
DB 1322 GCATGCTAGATGCTCAAGTGCATGACAGAGATTTTGAACCTTAATTTGCTCCGAGCCAG 1381  
QY 1382 ACATGCCAAAGGTACCTTTGTGATCGACTCTCCCAATTTTGTGATGATGAGCTGGGT 1441  
DB 1382 ACATGCCAAAGGTACCTTTGTGATCGACTCTCCCAATTTTGTGATGATGAGCTGGGT 1441

QY 1442 TAAAGTGTGCCAAGGAAGTGCATTGTCAATAGCAATAGCTGTGAAGGAGAGAGGACG 1501  
DB 1442 TAAAGTGTGCCAAGGAAGTGCATTGTCAATAGCAATAGCTGTGAAGGAGAGAGGACG 1501  
QY 1502 ACTTCTTGAGAAAGCCAGGAAGATTTAAAAGCTATGAGCTGCTATGTTGGTGTGCTGCTT 1561  
DB 1502 ACTTCTTGAGAAAGCCAGGAAGATTTAAAAGTATGAGCTGCTATGTTGGTGTGCTGCTT 1561  
QY 1562 TTGATGAAGAAGCAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCAACCCGGGCT 1621  
DB 1562 TTGATGAAGAAGCAGGCAACAGAAAACAGACACAAAATCAGAGTGTGCAACCCGGGCT 1621  
QY 1622 ACCATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA 1681  
DB 1622 ACCATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATTAATTTTGAACCTTAATA 1681  
QY 1682 TCCTAAACATTTGGGACTGGAATGGAGGAACAACTTGTATGCCATTAATTTTATCCATG 1741  
DB 1682 TCCTAAACATTTGGGACTGGAATGGAGGAACAACTTGTATGCCATTAATTTTATCCATG 1741  
QY 1742 CAAACAAAAGTCAATTAAGAAAACATTTACCTGGAGCCAGAAATAAGTGGAGTCTTTTCCAACT 1801  
DB 1742 CAAACAAAAGTCAATTAAGAAAACATTTACCTGGAGCCAGAAATAAGTGGAGTCTTTTCCAACT 1801  
QY 1802 TGTCTCTTCTTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGATGGGGTTTTTCCCTTT 1861  
DB 1802 TGTCTCTTCTTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGATGGGGTTTTTCCCTTT 1861  
QY 1862 ACCATGCAATCAAGTCTGCAATGGAGCATGGAGATAGTGAATGCTGGAACCTTCCCTGTCT 1921  
DB 1862 ACCATGCAATCAAGTCTGCAATGGAGCATGGAGATAGTGAATGCTGGAACCTTCCCTGTCT 1921  
QY 1922 ATGATGATATCCATAAGGAACTTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAGAGCC 1981  
DB 1922 ATGATGATATCCATAAGGAACTTTCTGCAGCTCTGTGAAGATCTCATCTGGAATAAGAGCC 1981  
QY 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2041  
DB 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2041  
QY 2042 TCATTCAGACTCATGAGTGGAGAAATGGCCCTGTGGAAGAGCCCTTGAAGTATGCCCTTG 2101  
DB 2042 TCATTCAGACTCATGAGTGGAGAAATGGCCCTGTGGAAGAGCCCTTGAAGTATGCCCTTG 2101  
QY 2102 TGAAGGGCAATTAAGAAAACATATTTATGAGGATCTTGAGGAGCCAGGTTAAACCAAAAA 2161  
DB 2102 TGAAGGGCAATTAAGAAAACATATTTATGAGGATCTTGAGGAGCCAGGTTAAACCAAAAA 2161  
QY 2162 AATATCCCGGACTCTCAATATAATTGAAGGAGCCCTGATGAATGGAATGGAATTTGTTG 2221  
DB 2162 AATATCCCGGACTCTCAATATAATTGAAGGAGCCCTGATGAATGGAATGGAATTTGTTG 2221  
QY 2222 GTGATCTTTTGGAGCTGGAAAATTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2281  
DB 2222 GTGATCTTTTGGAGCTGGAAAATTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2281  
QY 2282 TGAAGAAGGCTGTTGGCCACCTTATCCCTTTTATGGAAGGAGGAGGAGGAGGAGGAGGAG 2341  
DB 2282 TGAAGAAGGCTGTTGGCCACCTTATCCCTTTTATGGAAGGAGGAGGAGGAGGAGGAGGAG 2341  
QY 2342 TGTCTTAAACGGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2401  
DB 2342 TGTCTTAAACGGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2401  
QY 2402 TTAAGGGGAGCGTGACGACATAGGCAAGACATAGTTGGAGTACTCTTGGCTGCAATA 2461  
DB 2402 TTAAGGGGAGCGTGACGACATAGGCAAGACATAGTTGGAGTACTCTTGGCTGCAATA 2461  
QY 2462 ATTTCCGAGTATTTGATTTTAGGAGTCAAGTCTTCCATGTGATGAAGTACTGGAAGCTGCTC 2521  
DB 2462 ATTTCCGAGTATTTGATTTTAGGAGTCAAGTCTTCCATGTGATGAAGTACTGGAAGCTGCTC 2521  
QY 2522 TTGACCAACAAAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTCTTCCCTGGATGAAA 2581

```
Db 2522 TTGACCAAGAGGAGATATAATTTGGCTGTGAGGACTCATCACTCTTCCCTGGATGAAA 2581
Qy 2582 TGATTTTGTGTCAAAGGAAATGGAGAGATTAGCTATAAGGATTCOAATGTTGATGGAG 2641
Db 2582 TGATTTTGTGTCAAAGGAAATGGAGAGATTAGCTATAAGGATTCOAATGTTGATGGAG 2641
Qy 2642 GAGCAACCACTTCMAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701
Db 2642 GAGCAACCACTTCMAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701
Qy 2702 TAAATCAATGTCCTGAGCGGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761
Db 2702 TAAATCAATGTCCTGAGCGGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761
Qy 2762 ATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATAGACAGGACC 2821
Db 2762 ATCTAAAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATAGACAGGACC 2821
Qy 2822 ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAAGTCAAGCCAGAAAGTGTT 2881
Db 2822 ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAAGTCAAGCCAGAAAGTGTT 2881
Qy 2882 TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGACCCAGG 2941
Db 2882 TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGACCCAGG 2941
Qy 2942 TCTTTGAAGACTATGACCTGACAGAGCTGGTGAATACATGATGACGAAAGCCCTTTCTTTG 3001
Db 2942 TCTTTGAAGACTATGACCTGACAGAGCTGGTGAATACATGATGACGAAAGCCCTTTCTTTG 3001
Qy 3002 ATGTCTGGAGCTCCGGGGCAAGTACCCGNAATCGAGCTTCCCAAGATATTTAAGCACA 3061
Db 3002 ATGTCTGGAGCTCCGGGGCAAGTACCCGNAATCGAGCTTCCCAAGATATTTAAGCACA 3061
Qy 3062 AAACAGTAGGTGAGAGGCGAGAGAGGTCTACGATGATGCCCAACAATATGCTCAACACAC 3121
Db 3062 AAACAGTAGGTGAGAGGCGAGAGAGGTCTACGATGATGCCCAACAATATGCTCAACACAC 3121
Qy 3122 TGATTAAGTCAAAAGAACTCCGGGCGGGGTGTGGTTGGTTCTGGCCAGCACAGAGTA 3181
Db 3122 TGATTAAGTCAAAAGAACTCCGGGCGGGGTGTGGTTGGTTCTGGCCAGCACAGAGTA 3181
Qy 3182 TCCAAGACGACATTCACCTGTACGAGAGCTGCTGTGCCCGAGGCTGCAGAGCCCATAG 3241
Db 3182 TCCAAGACGACATTCACCTGTACGAGAGCTGCTGTGCCCGAGGCTGCAGAGCCCATAG 3241
Qy 3242 CCACCTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACGGAGCCATACT 3301
Db 3242 CCACCTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACGGAGCCATACT 3301
Qy 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCCTGGCATCCGTAATGATGGGCTGT 3361
Db 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCCTGGCATCCGTAATGATGGGCTGT 3361
Qy 3362 TTGCGGTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAG 3421
Db 3362 TTGCGGTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAG 3421
Qy 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCGAAG 3481
Db 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCGAAG 3481
Qy 3482 AGCTCCA TGAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541
Db 3482 AGCTCCA TGAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541
Qy 3542 TCGCAGACCTGCGAGGTTGCGGTACAAAGGCGATCCCGCGGCTCCTGGCTACCCAGCC 3601
Db 3542 TCGCAGACCTGCGAGGTTGCGGTACAAAGGCGATCCCGCGGCTCCTGGCTACCCAGCC 3601
Qy 3602 AGCCCGACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
Db 3602 AGCCCGACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
```

```
Db 3602 AGCCCGACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
Qy 3662 GCATTAAGTTAAACAGAAATCATTAGCAATGGACCTGCTTACGAGCTCTCAGGCTCTACT 3721
Db 3662 GCATTAAGTTAAACAGAAATCATTAGCAATGGACCTGCTTACGAGCTCTCAGGCTCTACT 3721
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAAGATTTCCAAAGGATCAGGTTGAGG 3781
Db 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAAGATTTCCAAAGGATCAGGTTGAGG 3781
Qy 3782 ATTATGCAATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841
Db 3782 ATTATGCAATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3875
Db 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3875

RESULT 3
AR482320 LOCUS 3856 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 75 from patent US 6703197.
ACCESSION AR482320
VERSION AR482320.1 GI:47244609
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3856)
AUTHORS Gravel,R.A., Rozen,R., LeClerc,D., Goyette,P. and Campeau,E.
TITLE Human methionine synthase: cloning, and methods for evaluating risk
of neural tube defects, cardiovascular disease, and cancer
JOURNAL Patent: US 6703197-A 75 09-MAR-2004;
Martinez R&D, Inc.; Montreal;
CAX;
FEATURES Location/Qualifiers
source 1..3856
/mol_type="genomic DNA"
ORIGIN
Query Match 91.9%; Score 3601; DB 6; Length 3856;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3851; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 64 ATGTCAACCCGCGCTCCAGACCTGTCCGAAACCCGAGAGGTCTGAAGAAAACCCCTGCGGGAT 123
Db 1 ATGTCAACCCGCGCTCCAGACCTGTCCGAAACCCGAGAGGTCTGAAGAAAACCCCTGCGGGAT 60
Qy 124 GAGATCAATGCCATTTCTGCAGAGAGGATTTATGGTCTGGATGGAGGGATGGGACCATG 183
Db 61 GAGATCAATGCCATTTCTGCAGAGAGGATTTATGGTCTGGATGGAGGGATGGGACCATG 120
Qy 184 ATCCAGCGGGAGAGGCTTAAACGAAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC 243
Db 121 ATCCAGCGGGAGAGGCTTAAACGAAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC 180
Qy 244 AGCCGCTGAAAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTTACCAA 303
Db 181 AGCCGCTGAAAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTTACCAA 240
Qy 304 ATCCATAAGGAATACATTTGCTGGCTGGGCGAGATATCATTTGAAACAAATACATTTTAGCAGC 363
Db 241 ATCCATAAGGAATACATTTGCTGGCTGGGCGAGATATCATTTGAAACAAATACATTTTAGCAGC 300
Qy 364 ACTAGTATTGCCCAAGCTGACTATGCTTTGAAACACTTGGCCTACCGGATGAACATGTGC 423
Db 301 ACTAGTATTGCCCAAGCTGACTATGCTTTGAAACACTTGGCCTACCGGATGAACATGTGC 360
Qy 424 TCTGCAGGAGTGGCCAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAAATTAAGAGG 483
Db 361 TCTGCAGGAGTGGCCAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAAATTAAGAGG 420
```

QY 484 TTTTGGCAGGGGCTCTGGGTCGCACTAATAAGACACTCTCTGTGTCCCACTCTGTGGAA 543  
DB 421 TTTTGGCAGGGGCTCTGGGTCGCACTAATAAGACACTCTCTGTGTCCCACTCTGTGGAA 480  
QY 544 AGGCCGATTAATAGGAACATCAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCC 603  
DB 481 AGGCCGATTAATAGGAACATCAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCC 540  
QY 604 AAAGGACTTCTCGATGCGGGGTGATATCTTACTCATTTGAAACTATTTTGTACTGCGC 663  
DB 541 AAAGGACTTCTCGATGCGGGGTGATATCTTACTCATTTGAAACTATTTTGTACTGCGC 600  
QY 664 AATGCCAAGCAGCCTTGTGTGCACTCCAAAATCTTTTTCAGAGAAATATGCTCCCGG 723  
DB 601 AATGCCAAGCAGCCTTGTGTGCACTCCAAAATCTTTTTCAGAGAAATATGCTCCCGG 660  
QY 724 CCTATCTTTATTTTCAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCCGACAGACA 783  
DB 661 CCTATCTTTATTTTCAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCCGACAGACA 720  
QY 784 GGAGAGGATTTGTCTATCAGCGTGTCTCATGGAGAACCACTCTGCAATTTGGAATTAATTTGT 843  
DB 721 GGAGAGGATTTGTCTATCAGCGTGTCTCATGGAGAACCACTCTGCAATTTGGAATTAATTTGT 780  
QY 844 GCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATTAATTTGGAATTAATTTGTAACAAGCC 903  
DB 781 GCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATTAATTTGGAATTAATTTGTAACAAGCC 840  
QY 904 TAGTCTCTGTATCCCAATGAGTCTTCCCAACACACCTTTTGGTGAATGATGAACG 963  
DB 841 TAGTCTCTGTATCCCAATGAGTCTTCCCAACACACCTTTTGGTGAATGATGAACG 900  
QY 964 CCTTCTATGATGCGCAAGCACTTAAAGGATTTTGTCTATGATGGCTTGTCAATATAGTT 1023  
DB 901 CCTTCTATGATGCGCAAGCACTTAAAGGATTTTGTCTATGATGGCTTGTCAATATAGTT 960  
QY 1024 GGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAGAGCTGTGAAAT 1083  
DB 961 GGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAGAGCTGTGAAAT 1020  
QY 1084 TGTAAAGCTAGCTTCCACTGCGCATCAACCACTTTTGTAAAGGACATATGCTGCTGCTA 1143  
DB 1021 TGTAAAGCTAGCTTCCACTGCGCATCAACCACTTTTGTAAAGGACATATGCTGCTGCTA 1080  
QY 1144 GAGCCCTTCAAGGATTTGACCGTACACCACTTTTGTAAAGGAGCGCTGTAATGTT 1203  
DB 1081 GAGCCCTTCAAGGATTTGACCGTACACCACTTTTGTAAAGGAGCGCTGTAATGTT 1140  
QY 1204 GAGGATCAAGGAATTTGTAAACTCATGCGCAGGAACATATGAAGAAGCCTTGTGT 1263  
DB 1141 GAGGATCAAGGAATTTGTAAACTCATGCGCAGGAACATATGAAGAAGCCTTGTGT 1200  
QY 1264 GTTCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATGC 1323  
DB 1201 GTTCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTGGATGTCAACATGGATGATGC 1260  
QY 1324 ATGCTAGATGGTCAAGTGCATGACCAAGATTTTGAACCTTAATTTGCTCCGAGCAGAC 1383  
DB 1261 ATGCTAGATGGTCAAGTGCATGACCAAGATTTTGAACCTTAATTTGCTCCGAGCAGAC 1320  
QY 1384 ATGCTAGATGGTCAAGTGCATGACCTCTCCAAATTTTGTCTGATTTGAAGCTGGGTTA 1443  
DB 1321 ATGCTAGATGGTCAAGTGCATGACCTCTCCAAATTTTGTCTGATTTGAAGCTGGGTTA 1380  
QY 1444 AAGTGTGCCAAGGAGTGCATTTGCAATAGCATTTAGTCTGAAGAGAGGAGGACGAC 1503  
DB 1381 AAGTGTGCCAAGGAGTGCATTTGCAATAGCATTTAGTCTGAAGAGAGGAGGACGAC 1440  
QY 1504 TTCTTGGAGAGCCAGGAGATTTAAAGTATGGAGCTGCTATGTTGTGCTATGGCTTTT 1563  
DB 1441 TTCTTGGAGAGCCAGGAGATTTAAAGTATGGAGCTGCTATGTTGTGCTATGGCTTTT 1500

QY 1564 GATGAGAGGACAGCGACAGAAACAGACACAAATAATCAGAGTGTGCAACCGGGCCTAC 1623  
DB 1501 GATGAGAGGACAGCGACAGAAACAGACACAAATAATCAGAGTGTGCAACCGGGCCTAC 1560  
QY 1624 CATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTTCACCCCTAATATC 1683  
DB 1561 CATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTTCACCCCTAATATC 1620  
QY 1684 CTAAACCATTTGGGACTTGAATGGAGGAAACAACCTTGTATGCAATTAATTTTATCCATGCA 1743  
DB 1621 CTAAACCATTTGGGACTTGAATGGAGGAAACAACCTTGTATGCAATTAATTTTATCCATGCA 1680  
QY 1744 ACAAAGTCAATTAAGAAACATTAACCTGAGGACCAATTAAGTGGAGTCTTTCCCACTTG 1803  
DB 1681 ACAAAGTCAATTAAGAAACATTAACCTGAGGACCAATTAAGTGGAGTCTTTCCCACTTG 1740  
QY 1804 TCCTTCTCTCCGAGGAATGGAACCAATTCAGAGCAATGCAATGGGGTTCCTTTTAC 1863  
DB 1741 TCCTTCTCTCCGAGGAATGGAACCAATTCAGAGCAATGCAATGGGGTTCCTTTTAC 1800  
QY 1864 CATGCAATCAAGTCTGGCATGGAACAATGAGATAGTGAATGCTGGAAACCTCCTCTGTGTAT 1923  
DB 1801 CATGCAATCAAGTCTGGCATGGAACAATGAGATAGTGAATGCTGGAAACCTCCTCTGTGTAT 1860  
QY 1924 GATGATATCCATAAGGAACCTTCTGCACTCTGTAAGATCTCATCTGGAAATAAGACCCCT 1983  
DB 1861 GATGATATCCATAAGGAACCTTCTGCACTCTGTAAGATCTCATCTGGAAATAAGACCCCT 1920  
QY 1984 GAGGCACATGAGAGCTCTTACGTTATGCCAGACTCAAGGACACAGAGGAGGAAAGATC 2043  
DB 1921 GAGGCACATGAGAGCTCTTACGTTATGCCAGACTCAAGGACACAGAGGAGGAAAGATC 1980  
QY 2044 ATTCAGACTGATGAGTGGAGAAATGSCCTGTGCGAAGAACGCTTGGATATGCTTGTG 2103  
DB 1981 ATTCAGACTGATGAGTGGAGAAATGSCCTGTGCGAAGAACGCTTGGATATGCTTGTG 2040  
QY 2104 AAGGGCATTTGAAAAACATTAATTTGAGGATCTAGGAGAACGAGTTAAACCAAAAAA 2163  
DB 2041 AAGGGCATTTGAAAAACATTAATTTGAGGATCTAGGAGAACGAGTTAAACCAAAAAA 2100  
QY 2164 TATCCCGACCTCTCAATTAATTTGAAGCAACCTCTGATGAATGGAATGAAATTTGTGTT 2223  
DB 2101 TATCCCGACCTCTCAATTAATTTGAAGCAACCTCTGATGAATGGAATGAAATTTGTGTT 2160  
QY 2224 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAGCCGCGGTATG 2283  
DB 2161 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAGCCGCGGTATG 2220  
QY 2284 AAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAACCAAGATG 2343  
DB 2221 AAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAACCAAGATG 2280  
QY 2344 CTTAAGCGCACAGTGAAGAGGACCTTACAGGGGACCATCTGCTGGCCACTGTT 2403  
DB 2281 CTTAAGCGCACAGTGAAGAGGACCTTACAGGGGACCATCTGCTGGCCACTGTT 2340  
QY 2404 AAAGGCGACGTGCACACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGCAATAAT 2463  
DB 2341 AAAGGCGACGTGCACACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGCAATAAT 2400  
QY 2464 TTCCGAGTATTTGATTTAGGAGTCAATGCTCAATGATGAATAGTGAAGAGTGTCTTT 2523  
DB 2401 TTCCGAGTATTTGATTTAGGAGTCAATGCTCAATGATGAATAGTGAAGAGTGTCTTT 2460  
QY 2524 GACCAAAAGCAGATATATTTGGCCCTGTCAGGACTCATCTCTTCCCTGGATGAAATG 2583  
DB 2461 GACCAAAAGCAGATATATTTGGCCCTGTCAGGACTCATCTCTTCCCTGGATGAAATG 2520  
QY 2584 ATTTTGTGGCCAAAGGAATGGAGAGATTTAGCTATAAGGATTTCCATTTGTTGATGGAGA 2643  
DB 2521 ATTTTGTGGCCAAAGGAATGGAGAGATTTAGCTATAAGGATTTCCATTTGTTGATGGAGA 2580  
QY 2644 GCAACCATTTCAAAAAACCCACACAGCAGTTTAAATAGTCTCCGAGATACAGTGCACCTGTA 2703

2581	Db		GCAACCACTTTCAAAAAACCCACACAGCAGTGTAAATAATAGTCCGAGATACAGTGCACCTGTN	2640
2704	Qy		ATCCATGTCCTGGACGGCTCCAAAGATGTGGTGGTGTTCCTCCAGCTGTTTAGATGAAAT	2763
2641	Db		NNCCATGTCCTGCACGGCTCAAGAGTGTGGTGGTGTTCCTCCAGCTGTTTAGATGAAAT	2700
2764	Qy		CTAAAGATGAATACTTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGACCAT	2823
2701	Db		CTAAAGATGAATACTTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGNCAT	2760
2824	Qy		TATGAGTCTCTCAAGGAGAGGAGATACTTACCCTTAAAGTCAAGCCAGAAAAAGTGTTTC	2883
2761	Db		TATGAGTCTCTCAAGGAGAGGAGATACTTACCCTTAAAGTCAAGCCAGAAAAAGTGTTTC	2820
2884	Qy		CAAAATGGAATGGCTGTCTGAACTCACCAGTGAAGCCACAGCTTATTGGGACCCAGTTC	2943
2821	Db		CAAAATGGAATGGCTGTCTGAACTCACCAGTGAAGCCACAGCTTATTGGGACCCAGTTC	2880
2944	Qy		TTTTGAAGACTATGACCTGCGAGAGCTGGTGGACTTACATTGACTTGGAAAGCCTTTCTTTGAT	3003
2881	Db		TTTTGAAGACTATGACCTGCGAGAGCTGGTGGACTTACATTGACTTGGAAAGCCTTTCTTTGAT	2940
3004	Qy		GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAGATATTATTAACGACAA	3063
2941	Db		GTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAGATATTATTAACGACAA	3000
3064	Qy		ACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCACACATATGCTGGAACACATG	3123
3001	Db		ACAGTAGGTGGAGAGGCCAGGAAGGTCTACGATGATGCCACACATATGCTGGAACACATG	3060
3124	Qy		ATTAGTCAAAAGAAACTCCCGGCCCGGGGTGTGGTTCGGCTTCTGGCCAGACACAGATATC	3183
3061	Db		ATTAGTCAAAAGAAACTCCCGGCCCGGGGTGTGGTTCGGCTTCTGGCCAGACACAGATATC	3120
3184	Qy		CAAGACGACATTCACCTGTATGCGAGAGGCTGCTGTGCTCCAGGCTCAGAGCCCATAGCC	3243
3121	Db		CAAGACGACATTCACCTGTATGCGAGAGGCTGCTGTGCTCCAGGCTCAGAGCCCATAGCC	3180
3244	Qy		ACTTTCTATGGGTTAAGCAACAGGCTGAGNAGACTCTGCCAGCAGGAGCCATACTAC	3303
3181	Db		ACTTTCTATGGGTTAAGCAACAGGCTGAGNAGACTCTGCCAGCAGGAGCCATACTAC	3240
3304	Qy		TGCCTCTCAGACTTCATCGCTCCCTTCGATTCCTGGCATCCGTGACTACTCTGGGCTGTTT	3363
3241	Db		TGCCTCTCAGACTTCATCGCTCCCTTCGATTCCTGGCATCCGTGACTACTCTGGGCTGTTT	3300
3364	Qy		GCCGTTGCTCTTTGGGGTAGAAGAGCTGAGCAAGGCCTATGAGGATGATGGTGACGAC	3423
3301	Db		GCCGTTGCTCTTTGGGGTAGAAGAGCTGAGCAAGGCCTATGAGGATGATGGTGACGAC	3360
3424	Qy		TACAGCAGCATCATGTTCAAGCGCTGGGGAGCCGGCTGGCAGAGGCTTTTGCAGAGAG	3483
3361	Db		TACAGCAGCATCATGTTCAAGCGCTGGGGAGCCGGCTGGCAGAGGCTTTTGCAGAGAG	3420
3484	Qy		CTCCATGAAGAGTTCCGCCGAGAACTGTGGCCTACTGTGCGACTGAGCAGCTGAGCTC	3543
3421	Db		CTCCATGAAGAGTTCCGCCGAGAACTGTGGGCCCTACTGTGCGACTGAGCAGCTGAGCTC	3480
3544	Qy		GCAGACTTCGAAGGTTTCGGGTACAAAGGGCAATCCGCCCGGCTCCTGGCTACCCACAGCCAG	3603
3481	Db		GCAGACTTCGAAGGTTTCGGGTACAAAGGGCAATCCGCCCGGCTCCTGGCTACCCACAGCCAG	3540
3604	Qy		CCCGACCAACCGGAAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC	3663
3541	Db		CCCGACCAACCGGAAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC	3600
3664	Qy		ATTAGGTTAAACAGATCATTTAGCAATGGCACTGCTTCAGCAGTCTCAGGCTCTACTTC	3723
3601	Db		ATTAGGTTAAACAGATCATTTAGCAATGGCACTGCTTCAGCAGTCTCAGGCTCTACTTC	3660
3724	Qy		TCCAAATTTGAAGTCCAAAATATTTTGCTGTGGGGAGATTTTCCAAAGGATCAGGTTGAGGAT	3783

Db	3661	TCCAATTGTAAGTCCAAATAATTTTGTGTGGGGAAGATTTCCAGAGATCAGGTTGAGGAT	3722
Qy	3784	TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTTGAGAAATGGCTTGGACCACTTTTG	3843
Db	3721	TATGCATTGAGGAAGAACATATCTGTGGCTGAGGTTTGAGAAATGGCTTGGACCACTTTTG	3780
Qy	3844	GGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGGCCTTTTTTATCTTGATGATCCTCA	3903
Db	3781	GGATATGATACAGACTAACTTTTTTTTTTTTTTTTTTTTGGCCTTTTTTATCTTGATGATCCTCA	3840
Qy	3904	AGGAATACAACTAG 3919	
Db	3841	AGGAATACAACTAG 3856	
RESULT 4			
AR144957			
LOCUS	AR144957	7122 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 2 from patent US 6210950.		
ACCESSION	AR144957		
VERSION	AR144957.1	GI:15106824	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 7122)		
TITLE	Johnson, W.G. and Stenroos, E. Scott.		
JOURNAL	Methods for diagnosing, preventing, and treating developmental		
FEATURES	disorders due to a combination of genetic and environmental factors		
source	Patent: US 6210950-A 2 03-APR-2001;		
	Location/Qualifiers		
	1. .7122		
	/organism="unknown"		
ORIGIN	/mol_type="unassigned DNA"		
Query Match	89.74%;	Score 3517;	DB 6; Length 7122;
Best Local Similarity	99.88%;	Pred. No. 0;	
Matches 3867;	Conservative 0;	Mismatches 7;	Indels 0; Gaps 0;
Qy	2	GTCACTGTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	61
Db	225	GTCACTGTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA	284
Qy	62	ACATGTCAACCGCGCTCAAGACTGTGCGCAACCCGGAAGTCTGAAAGAAACCTCGCGG	121
Db	285	ACATGTCAACCGCGCTCAAGACTGTGCGCAACCCGGAAGTCTGAAAGAAACCTCGCGG	344
Qy	122	ATGAGATCAATGCCATTCTCGAAGAGGATTATGTGTCTGGATGGAGGATGGGACCA	181
Db	345	ATGAGATCAATGCCATTCTCGAAGAGGATTATGTGTCTGGATGGAGGATGGGACCA	404
Qy	182	TGATCCAGCGGAGAGACTAAACGAAGAACATCTTCGAGGTGAGGAATTTAAAGATCATG	241
Db	405	TGATCCAGCGGAGAGACTAAACGAAGAACATCTTCGAGGTGAGGAATTTAAAGATCATG	464
Qy	242	CCAGGCGCTGAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCATTTTACC	301
Db	465	CCAGGCGCTGAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCATTTTACC	524
Qy	302	AAATCCATAAGGAATATCTTGCTGCTGGGGCAGATATCATTTGAAACAAATACTTTTACGA	361
Db	525	AAATCCATAAGGAATATCTTGCTGCTGGGGCAGATATCATTTGAAACAAATACTTTTACGA	584
Qy	362	GCACTAGTATTTGCCCAAGCTGACTATGGCCCTTGAACACTTGGCCCTACGGGATGAACATGT	421
Db	585	GCACTAGTATTTGCCCAAGCTGACTATGGCCCTTGAACACTTGGCCCTACGGGATGAACATGT	644
Qy	422	GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA	481
Db	645	GCTCTGCAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA	704
Qy	482	GGTTTGTGGCAGGGGCTCTGGGTCGGAATAAAGACACTCTCTGTGTGCCCATCTGTGG	541

Db 705 GGTGTTGGCAGGGCTCTGGGTCCGACTAATAAGACACTCTCTGTGTCCCACTGTGTG 764  
Qy 542 AAAGGCGGATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
Db 765 AAAGGCGGATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 824  
Qy 602 CCAAGGACTCTTGGATGGGGGGTGGATCTTACTCATTTGAACTATTTTGGATGACTG 661  
Db 825 CCAAGGACTCTTGGATGGGGGGTGGATCTTACTCATTTGAACTATTTTGGATGACTG 884  
Qy 662 CCAATGCAAGGAGCCTCTTTGCACTCCAAATCTTTTGAAGGAGAAATATGCTCCCC 721  
Db 885 CCAATGCAAGGAGCCTCTTTGCACTCCAAATCTTTTGAAGGAGAAATATGCTCCCC 944  
Qy 722 GGCCTATCTTTATTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGGACAGA 781  
Db 945 GGCCTATCTTTATTTCAAGGACGATCGTTGATAAAAGTGGCGGACTCTTTCCGGACAGA 1004  
Qy 782 CAGGAGGGATTTGTTCATCAGCGTGTCTCATCGAGAACCACTCTGCAATGGATTTAAAT 841  
Db 1005 CAGGAGGGATTTGTTCATCAGCGTGTCTCATCGAGAACCACTCTCAATGGATTTAAAT 1064  
Qy 842 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTCGAAATAATTTGAAATGTACACAG 901  
Db 1065 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTCGAAATAATTTGAAATGTACACAG 1124  
Qy 902 CCTATGCTCTGTATTCCTCAATGCAAGCTCTCCCAACACTTTGGTGACTATGATGAAA 961  
Db 1125 CCTATGCTCTGTATTCCTCAATGCAAGCTCTCCCAACACTTTGGTGACTATGATGAAA 1184  
Qy 962 CGCCTTCTATGATGGCCAGCACTAAAGATTTTGTCTATGAGTGGTGGTCAATATAG 1021  
Db 1185 CGCCTTCTATGATGGCCAGCACTAAAGATTTTGTCTATGAGTGGTGGTCAATATAG 1244  
Qy 1022 TTGAGGATGCTGTGGGTCAACACCAATCATATTCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1245 TTGAGGATGCTGTGGGTCAACACCAATCATATTCAGGGAATTTGCTGAAGCTGTGAAA 1304  
Qy 1082 ATTGTAAAGCTAGAGTTTCACTGCACTGCTTTTGAAGGACATATGTTACTGTGGTTC 1141  
Db 1305 ATTGTAAAGCTAGAGTTTCACTGCACTGCTTTTGAAGGACATATGTTACTGTGGTTC 1364  
Qy 1142 TAGAGCCCTTCAAGTTGACCCGTACACCAACTTTGTTAACTTGGAGAGCGCTGAATG 1201  
Db 1365 TAGAGCCCTTCAAGTTGACCCGTACACCAACTTTGTTAACTTGGAGAGCGCTGAATG 1424  
Qy 1202 TTGAGGATCAAGGATTTGCTTAACTCATATGCGAGGAACTATGAAGACCTTGT 1261  
Db 1425 TTGAGGATCAAGGATTTGCTTAACTCATATGCGAGGAACTATGAAGACCTTGT 1484  
Qy 1262 GTGTGCAAGTGCAGGTGGAATGGAGCCAGGTGTTGGATGTCAACATGGAATGATG 1321  
Db 1485 GTGTGCAAGTGCAGGTGGAATGGAGCCAGGTGTTGGATGTCAACATGGAATGATG 1544  
Qy 1322 GCATGCTAGATGTTCAAGTGCATGACCAAGTTTTCGCACTTAAATGCTTCCGAGCCAG 1381  
Db 1545 GCATGCTAGATGTTCAAGTGCATGACCAAGTTTTCGCACTTAAATGCTTCCGAGCCAG 1604  
Qy 1382 ACATGCAAGGATACCTTTGTGATGACCTCTCCAAATTTGCTGTGATTTGAAGCTGGGT 1441  
Db 1605 ACATGCAAGGATACCTTTGTGATGACCTCTCCAAATTTGCTGTGATTTGAAGCTGGGT 1664  
Qy 1442 TAAAGTGTCCCAAGGAGTGCATTTGTAATAGCATTTAGTCTGAAGGAGGAGGAGCG 1501  
Db 1665 TAAAGTGTCCCAAGGAGTGCATTTGTAATAGCATTTAGTCTGAAGGAGGAGGAGCG 1724  
Qy 1502 ACTTCTTGGAGAGGCCAGGAAGTTTAAAGATGATGAGCTGTATGTTGGTCAATGGCTT 1561  
Db 1725 ACTTCTTGGAGAGGCCAGGAAGTTTAAAGATGATGAGCTGTATGTTGGTCAATGGCTT 1784  
Qy 1562 TTGATGAAGGAGGAGGCAAGAAACAGACACAAATTCAGATGTGCAACCCGGGCT 1621

Db 1785 TTGATGAAGAGCAGGGCAACAGAAACAGACACAAAAATCAGAGTGTGCAACCCGGGCT 1844  
Qy 1622 ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGGACCTTAATA 1681  
Db 1845 ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGGACCTTAATA 1904  
Qy 1682 TCCTAACCATTTGGGACTGGAATGGAGGAACAACTTGTGTATGCCATTATTTTATCATG 1741  
Db 1905 TCCTAACCATTTGGGACTGGAATGGAGGAACAACTTGTGTATGCCATTATTTTATCATG 1964  
Qy 1742 CAAACAAAGTCATTAAGAAACATTTACCTGGAGCCAGAAATAAGTGAGAGTCTTTTCCAAT 1801  
Db 1965 CAAACAAAGTCATTAAGAAACATTTACCTGGAGCCAGAAATAAGTGAGAGTCTTTTCCAAT 2024  
Qy 1802 TGTCTCTTCTTCCGAGGAATGGAAGCAATTCGAGAAACAAATGCAATGGGTTTTTCTTT 1861  
Db 2025 TGTCTCTTCTTCCGAGGAATGGAAGCAATTCGAGAAACAAATGCAATGGGTTTTTCTTT 2084  
Qy 1862 ACCATGCAATCAAGTCTGCAATGGAGATGTAATGCTGGAACCTTCCCTGTGT 1921  
Db 2085 ACCATGCAATCAAGTCTGCAATGGAGATGTAATGCTGGAACCTTCCCTGTGT 2144  
Qy 1922 ATGATGATATCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 1981  
Db 2145 ATGATGATATCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 2204  
Qy 1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCCAGACTCAAGGCCACAGGAGGAAGAAAG 2041  
Db 2205 CTGAGGCCACTGAGAGCTCTTACGTTATGCCCAGACTCAAGGCCACAGGAGGAAGAAAG 2264  
Qy 2042 TCATTGAGCTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGAATGATGCCCTTG 2101  
Db 2265 TCATTGAGCTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGAATGATGCCCTTG 2324  
Qy 2102 TGAAGGCCATTGAAAAACATATTTTGAAGTACTGAGGAAGCCAGTTTAAACCAAAAA 2161  
Db 2325 TGAAGGCCATTGAAAAACATATTTTGAAGTACTGAGGAAGCCAGTTTAAACCAAAAA 2384  
Qy 2162 AATATCCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGAATGAAATTTGTTG 2221  
Db 2385 AATATCCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGAATGAAATTTGTTG 2444  
Qy 2222 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGGTATTAAGTCAAGCCCGGTTA 2281  
Db 2445 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGGTATTAAGTCAAGCCCGGTTA 2504  
Qy 2282 TGAAGAGGCTCTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAACCCAGAG 2341  
Db 2505 TGAAGAGGCTCTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAACCCAGAG 2564  
Qy 2342 TGCCTTAAACGACAGTAGAAGAGGAGCCCTTACAGGGCCACCATCTGCTGGCCACTG 2401  
Db 2565 TGCCTTAAACGACAGTAGAAGAGGAGCCCTTACAGGGCCACCATCTGCTGGCCACTG 2624  
Qy 2402 TTAAAGGCGACGTGCAACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGGCTGCAATA 2461  
Db 2625 TTAAAGGCGACGTGCAACGACATAGGCAAGAAACATAGTTGGAGTAGTCTTGGCTGCAATA 2684  
Qy 2462 ATTTCCGAGTTATGATTTAGGAGTCACTGCTCAATGATTAAGTACTGGAAGCTGCTC 2521  
Db 2685 ATTTCCGAGTTATGATTTAGGAGTCACTGCTCAATGATTAAGTACTGGAAGCTGCTC 2744  
Qy 2522 TTGACCAACAAAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTGGATGAAA 2581  
Db 2745 TTGACCAACAAAGCAGATATAATTTGGCCTGTGAGGACTCATCTCTTCCCTGGATGAAA 2804  
Qy 2582 TGAATTTTGTGCCAAGGAAATGAGAGATTAAGCTATTAAGGATTCATTTGTTGATGGAG 2641  
Db 2805 TGAATTTTGTGCCAAGGAAATGAGAGATTAAGCTATTAAGGATTCATTTGTTGATGGAG 2864  
Qy 2642 GAGCAACCACTTCAAAAACCCACAGCAGGTTTAAATAGCTCGAGATACAGTGCACTG 2701  
Db 2865 GAGCAACCACTTCAAAAACCCACAGCAGGTTTAAATAGCTCGAGATACAGTGCACTG 2924







QY 602 CCAAAGGACTTCTGGATGGGGGTTGATATCTTATCTCATGTGAAACTATTTTTGATGACTG 661  
DB 825 CCAAAGGACTTCTGGATGGGGGTTGATATCTTACTCATTTGAAACTATTTTTGATGACTG 884  
QY 662 CCAATGCCAAGGAGCCTCTGTTGCACTCCAAATCTTTTGGAGGAATATGCTCCCC 721  
DB 885 CCAATGCCAAGGAGCCTCTGTTGCACTCCAAATCTTTTGGAGGAATATGCTCCCC 944  
QY 722 GGCTATCTTTTATTTTCAGGACAGATCGTTGATAAAGTGGCGGACTCTTTTCGGGACAGA 781  
DB 945 GGCTATCTTTTATTTTCAGGACAGATCGTTGATAAAGTGGCGGACTCTTTTCGGGACAGA 1004  
QY 782 CAGGAGGAGATTGTTCATCAGCGTGTCTCATGGAGAACCACTCTCTGCAATTTGGATTTAAAT 841  
DB 1005 CAGGAGGAGATTGTTCATCAGCGTGTCTCATGGAGAACCACTCTCATTTGGATTTAAAT 1064  
QY 842 GTCTTTGGGTGCGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATGTACACAG 901  
DB 1065 GTCTTTGGGTGCGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATGTACACAG 1124  
QY 902 CCTATGCTCTCTGTTATCCCAATGCGAGTCTTCCCAACACTTTGGTGACTATGATGAAA 961  
DB 1125 CCTATGCTCTCTGTTATCCCAATGCGAGTCTTCCCAACACTTTGGTGACTATGATGAAA 1184  
QY 962 CGCTTCTATGATGGCCAAAGCACTAAAGATTTTGCTATGGATGGCTTGGTCAATATAG 1021  
DB 1185 CGCTTCTATGATGGCCAAAGCACTAAAGATTTTGCTATGGATGGCTTGGTCAATATAG 1244  
QY 1022 TTGGAGATGCTCTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
DB 1245 TTGGAGATGCTCTGGGTCAACACCAAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTCTTTTGAAGACATATGTTACTCTGGTC 1141  
DB 1305 ATTGTAAAGCTTAGAGTTCCACCTGCCACTCTTTTGAAGACATATGTTACTCTGGTC 1364  
QY 1142 TAGAGCCTTTCAGGATTTGACCGGTACACCAACTTTGTTTAACTATGGAGAGCGCTGTAATG 1201  
DB 1365 TAGAGCCTTTCAGGATTTGACCGGTACACCAACTTTGTTTAACTATGGAGAGCGCTGTAATG 1424  
QY 1202 TTGCAGGATCAAGGATTTGCTTAACTCATATGCGAGGAACCTATGAAGACCTTGT 1261  
DB 1425 TTGCAGGATCAAGGATTTGCTTAACTCATATGCGAGGAACCTATGAAGAGCCTTGT 1484  
QY 1262 GTGTTCCAAAGTGCAGTGCAGTAATGGAGCCCAAGTGTGGATGTCAACATCGATGATG 1321  
DB 1485 GTGTTCCAAAGTGCAGTGCAGTAATGGAGCCCAAGTGTGGATGTCAACATCGATGATG 1544  
QY 1322 GCATGCTAGATGTTCCAAAGTGCAGTAATGGAGCCCAAGTGTGGATGTCAACATCGATG 1381  
DB 1545 GCATGCTAGATGTTCCAAAGTGCAGTAATGGAGCCCAAGTGTGGATGTCAACATCGATG 1604  
QY 1382 ACATCGAAAGGTACCTTTGTGATGCACTCTCCCAATTTTGTCTGTGATTTGAAGCTGGGT 1441  
DB 1605 ACATCGAAAGGTACCTTTGTGATGCACTCTCCCAATTTTGTCTGTGATTTGAAGCTGGGT 1664  
QY 1442 TAAAGTGTCTCCAAAGGAGTGCATTTGCTCAATAGCAATTTAGTCTGAAGGAGGAGAGACG 1501  
DB 1665 TAAAGTGTCTCCAAAGGAGTGCATTTGCTCAATAGCAATTTAGTCTGAAGGAGGAGAGACG 1724  
QY 1502 ACTTCTTGGAGAGGCGCAGGAAGATTAAAGATGATGAGCTGCTATGCTGTGGTCAATGGCTT 1561  
DB 1725 ACTTCTTGGAGAGGCGCAGGAAGATTAAAGATGATGAGCTGCTATGCTGTGGTCAATGGCTT 1784  
QY 1562 TTGATGAAGAGGAGGCGCAACAGAAACAGACACAAATAATCAGAGTGTGACCCGGGCT 1621  
DB 1785 TTGATGAAGAGGAGGCGCAACAGAAACAGACACAAATAATCAGAGTGTGACCCGGGCT 1844  
QY 1622 ACCATCTGCTGTGAAATACTGGCTTTTAACTCCAAATGACATTTTGTGACCCCTAATA 1681  
DB 1845 ACCATCTGCTGTGAAATACTGGCTTTTAACTCCAAATGACATTTTGTGACCCCTAATA 1904

QY 1682 TCCTAACCAATGGGACTGGAATGGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 1741  
DB 1905 TCCTAACCAATGGGACTGGAATGGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 1964  
QY 1742 CAAACAAAGCTCATTTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGAGGCTTTTCCAACT 1801  
DB 1965 CAAACAAAGCTCATTTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGAGGCTTTTCCAACT 2024  
QY 1802 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTGGAAGCAATGCAATGGGTTTTTCTCTTT 1861  
DB 2025 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTGGAAGCAATGCAATGGGTTTTTCTCTTT 2084  
QY 1862 ACATGCAATCAAGCTCTGCAATGGAATGAGTAGTAGTAATGCTGAAACCTCCCTGTGT 1921  
DB 2085 ACCATGCAATCAAGCTCTGCAATGGAATGAGTAGTAGTAATGCTGAAACCTCCCTGTGT 2144  
QY 1922 ATGATGATATCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTCGAATAAAGACC 1981  
DB 2145 ATGATGATATCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTCGAATAAAGACC 2204  
QY 1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAACAGGAGGGAAGAAAG 2041  
DB 2205 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAACAGGAGGGAAGAAAG 2264  
QY 2042 TCATTCAGACTGATGAGTGAGAAATGGCCCTGTGCAAGAACGCTTGTAGTATGCCCTTG 2101  
DB 2265 TCATTCAGACTGATGAGTGAGAAATGGCCCTGTGCAAGAACGCTTGTAGTATGCCCTTG 2324  
QY 2102 TGAAGGCCATTCGAAACCAATATTTATGAGGACTCTGAGGAGCCAGGTTAAACCAAAAAA 2161  
DB 2325 TGAAGGCCATTCGAAACCAATATTTATGAGGACTCTGAGGAGCCAGGTTAAACCAAAAAA 2384  
QY 2162 AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGAAATTTGTTG 2221  
DB 2385 AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGAAATTTGTTG 2444  
QY 2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTA 2281  
DB 2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAAGCCCGGTTA 2504  
QY 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACCAAGAG 2341  
DB 2505 TGAAGAGGCTGTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAGAAACCAAGAG 2364  
QY 2342 TGCTTAAACGGCAGTAGTAAGAGAGGACCCCTTACCAAGGCAACCTGCTGCGCCACTG 2401  
DB 2565 TGCTTAAACGGCAGTAGTAAGAGAGGACCCCTTACCAAGGCAACCTGCTGCGCCACTG 2624  
QY 2402 TTAAGGGCGAGCTGCAAGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2461  
DB 2625 TTAAGGGCGAGCTGCAAGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2684  
QY 2462 ATTTCCGAGTATTTAGGAGTATGAGTCCATGCTCAATAGATACCTGGAAGCTGCTC 2521  
DB 2685 ATTTCCGAGTATTTAGGAGTATGAGTCCATGCTCAATAGATACCTGGAAGCTGCTC 2744  
QY 2522 TTGACCAACAAAGCAGATATAATTTGGCCTCTCAGGACTCATCTCTCTCCCTGGATGAAA 2581  
DB 2745 TTGACCAACAAAGCAGATATAATTTGGCCTCTCAGGACTCATCTCTCTCCCTGGATGAAA 2804  
QY 2582 TGAATTTTGTTCGCAAGGAAATGGAGAGATTAGCTATTAAGGATTCATTTGTTGATTTGGAG 2641  
DB 2805 TGAATTTTGTTCGCAAGGAAATGGAGAGATTAGCTATTAAGGATTCATTTGTTGATTTGGAG 2864  
QY 2642 GAGCAACCACTTCGCAAGGAAACCCACAGCAGGTTAAATAGCTCCGAGTACAGTGCACCTG 2701  
DB 2865 GAGCAACCACTTCGCAAGGAAACCCACAGCAGGTTAAATAGCTCCGAGTACAGTGCACCTG 2924  
QY 2702 TAATCCATGCTCTGGAGCGGTCCAAGAGTGTGGTGTGTTCCAGCTGTTTAGATGAAA 2761  
DB 2925 TAATCCATGCTCTGGAGCGGTCCAAGAGTGTGGTGTGTTCCAGCTGTTTAGATGAAA 2984  
QY 2762 ATCTAAAGGATGAATATCTTTTGGAGGAATCATGGAGGAATATAGAGATATTTAGACAGACC 2821



|||||  
825 CCAAGACCTCTGGATGGGGGTTGATATCTTACTCATTTGAAACTATTTTTTGATCTG 884  
QY  
662 CCAATGCAAGGACGCTCTGTTTGCACTCCAAATCTTTTTTGAGGAGAAATATGCTCCCC 721  
Db  
885 CCAATGCCAAGGACGCTCTGTTTGCACTCCAAATCTTTTTTGAGGAGAAATATGCTCCCC 944  
QY  
722 GGCTATCTTTATTTCHAGGACAGATGTTGATATAAAGTGGCGGACTCTTTTCGGACAGA 781  
Db  
945 GGCTATCTTTATTTCHAGGACAGATGTTGATATAAAGTGGCGGACTCTTTTCGGACAGA 1004  
QY  
782 CAGGAGAGGATTTGTCATCAGCGTGTCTCATGAGAACACATCTGCAATTTGGATTAAT 841  
Db  
1005 CAGGAGAGGATTTGTCATCAGCGTGTCTCATGAGAACACATCTGCAATTTGGATTAAT 1064  
QY  
842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATATGTAACACAG 901  
Db  
1065 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATATGTAACACAG 1124  
QY  
902 CCTATGCTCTGTTATCCCAATGACGCTCTTCCCAACACCTTTGGTGACTATGATGAAA 961  
Db  
1125 CCTATGCTCTGTTATCCCAATGACGCTCTTCCCAACACCTTTGGTGACTATGATGAAA 1184  
QY  
962 CGCTTCTATGATGGCCCAAGCACTTAAGGATTTGCTATGATGCTGCTGCTCAATATAG 1021  
Db  
1185 CGCTTCTATGATGGCCCAAGCACTTAAGGATTTGCTATGATGCTGCTGCTCAATATAG 1244  
QY  
1022 TTGGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1081  
Db  
1245 TTGGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1304  
QY  
1082 ATTGTAGCTAGAGTTCCACTGCCACTGCTTTTGAAGACATATGTTACTGCTGTGTC 1141  
Db  
1305 ATTGTAGCTAGAGTTCCACTGCCACTGCTTTTGAAGACATATGTTACTGCTGTGTC 1364  
QY  
1142 TAGAGCCCTTCAGATTTGACCGGTACACCACTTTGTTTACATTTGAGAGCGCTGTAAATG 1201  
Db  
1365 TAGAGCCCTTCAGATTTGACCGGTACACCACTTTGTTTACATTTGAGAGCGCTGTAAATG 1424  
QY  
1202 TTCAGGATCAAGGAAGTTGCTTAACTCATATGCGAGGAAACTATGAAGAGCCCTGT 1261  
Db  
1425 TTCAGGATCAAGGAAGTTGCTTAACTCATATGCGAGGAAACTATGAAGAGCCCTGT 1484  
QY  
1262 GTGTTGCCAAGTGCAGTGGAAATGGAGCCAGCGTGTGATGTCACATGATGATG 1321  
Db  
1485 GTGTTGCCAAGTGCAGTGGAAATGGAGCCAGCGTGTGATGTCACATGATGATG 1544  
QY  
1322 GCATGCTAGATGTCGAAGTGCATGACAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1381  
Db  
1545 GCATGCTAGATGTCGAAGTGCATGACAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1504  
QY  
1382 ACATCGCAAGGTACCTTTGTGATCGACTCCTCCAAATTTGCTGTGATTTGAAGCTGGGT 1441  
Db  
1605 ACATCGCAAGGTACCTTTGTGATCGACTCCTCCAAATTTGCTGTGATTTGAAGCTGGGT 1664  
QY  
1442 TAAAGTCTGCCAAGGGAAGTGCATTTGCAATAGCATTAGTCTGCAAGGAGGAGAGACG 1501  
Db  
1665 TAAAGTCTGCCAAGGGAAGTGCATTTGCAATAGCATTAGTCTGCAAGGAGGAGAGACG 1724  
QY  
1502 ACTTCTTGGAGAGGCGCAGGAAGATTAAAGTATGAGCTGCTATGATGCTGCTCATGCTT 1561  
Db  
1725 ACTTCTTGGAGAGGCGCAGGAAGATTAAAGTATGAGCTGCTATGATGCTGCTCATGCTT 1784  
QY  
1562 TTGATGAAGAGGACGCGCAACAGAACAGACACAAATCAGAGTGTGACCCGGGCT 1621  
Db  
1785 TTGATGAAGAGGACGCGCAACAGAACAGACACAAATCAGAGTGTGACCCGGGCT 1844  
QY  
1622 ACCATCTGCTTGTGAAAAAATCTGGCTTTAATCCAAATGACATTATTTTGAACCTAATA 1681  
Db  
1845 ACCATCTGCTTGTGAAAAAATCTGGCTTTAATCCAAATGACATTATTTTGAACCTAATA 1904  
QY  
1682 TCCTAACCATTTGGACTGGAAATGGAGGAAACAACTTTGATGCCATTAAATTTATCCATG 1741  
Db  
|||||

Db  
1905 TCCTAACCATTTGGACTGGAAATGGAGGAAACAACTTTGTATGCCATTAAATTTATCCATG 1964  
QY  
1742 CAAACAAAGCTCATTAAGAAACATTAACCTGGAGCCAGAAATAAGTGAGGTCTTTTCAAAC 1801  
Db  
1965 CAAACAAAGCTCATTAAGAAACATTAACCTGGAGCCAGAAATAAGTGAGGTCTTTTCAAAC 2024  
QY  
1802 TGTCTCTCTCTCCGAGGAAATGGAGCCATTCGAGAACCAATGCAATGGGTTTCTCTTT 1861  
Db  
2025 TGTCTCTCTCTCCGAGGAAATGGAGCCATTCGAGAACCAATGCAATGGGTTTCTCTTT 2084  
QY  
1862 ACCATGCAATCAAGTCTGCGATGGACATGGAGATAGTGAATCTGGAACCTCCCTGTGT 1921  
Db  
2085 ACCATGCAATCAAGTCTGCGATGGACATGGAGATAGTGAATCTGGAACCTCCCTGTGT 2144  
QY  
1922 ATGATGATATCATTAAGGAACTTTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db  
2145 ATGATGATATCATTAAGGAACTTTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2204  
QY  
1982 CTGAGGCCACTGAGAGCTTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2041  
Db  
2205 CTGAGGCCACTGAGAGCTTTACGTTATGCCAGACTCAAGGCACAGAGGGAAGAAAG 2264  
QY  
2042 TCATTCCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG 2101  
Db  
2265 TCATTCCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG 2324  
QY  
2102 TGAAGGGCACTTGAACAACTATTTATTGAGGATACTGAGGAAGCCAGTTTAAACCAAAAA 2161  
Db  
2325 TGAAGGGCACTTGAACAACTATTTATTGAGGATACTGAGGAAGCCAGTTTAAACCAAAAA 2384  
QY  
2162 AATATCCCGACCTCTCAATATTAATTGAAGGACCCCTGATGAATGGAATGGAATTTGTTG 2221  
Db  
2385 AATATCCCGACCTCTCAATATTAATTGAAGGACCCCTGATGAATGGAATGGAATTTGTTG 2444  
QY  
2222 GTGATCTTTTGGAGCTGGAACAACTGTTCTTCTACCTCAGGTTTAAAGTTCAGCCCGGTTA 2281  
Db  
2445 GTGATCTTTTGGAGCTGGAACAACTGTTCTTCTACCTCAGGTTTAAAGTTCAGCCCGGTTA 2504  
QY  
2282 TGAAGAGGCTGTGTGGCCACCTTTATCCCTTTTCAATGGAACAAAGAGAGAAACACAGAG 2341  
Db  
2505 TGAAGAGGCTGTGTGGCCACCTTTATCCCTTTTCAATGGAACAAAGAGAGAAACACAGAG 2564  
QY  
2342 TGTTHAACGGCACAGTAGAAGAGAGGACCTTTACAGGGCACCATCTGCTGGCCACTG 2401  
Db  
2565 TGTTHAACGGCACAGTAGAAGAGAGGACCTTTACAGGGCACCATCTGCTGGCCACTG 2624  
QY  
2402 TTAAGGGCACGTTGACGACATAGGCAAGAACATAGTTCGAGTAGTCTTGGCTGCAATA 2461  
Db  
2625 TTAAGGGCACGTTGACGACATAGGCAAGAACATAGTTCGAGTAGTCTTGGCTGCAATA 2584  
QY  
2462 ATTTCCGAGTTATTTAGGAGTCAATGACTCCATGTGATAGATATCGAAAGCTGCTC 2521  
Db  
2685 ATTTCCGAGTTATTTAGGAGTCAATGACTCCATGTGATAGATATCGAAAGCTGCTC 2744  
QY  
2522 TTGACCAACAAAGCAGATATAATTTGGCTCTCAGGACTCATCACTCTTCCCTGGATGAAA 2581  
Db  
2745 TTGACCAACAAAGCAGATATAATTTGGCTCTCAGGACTCATCACTCTTCCCTGGATGAAA 2804  
QY  
2582 TGAATTTTCTTCCCAAGGAAATGGAGAGATTAGCTATAGGATTCCATTTGTTGATTTGGAG 2641  
Db  
2805 TGAATTTTCTTCCCAAGGAAATGGAGAGATTAGCTATAGGATTCCATTTGTTGATTTGGAG 2864  
QY  
2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTGTAAATAGCTCCGAGATACAGTGACCTG 2701  
Db  
2865 GAGCAACCACTTCAAAAAACCCACACAGCAGTGTAAATAGCTCCGAGATACAGTGACCTG 2924  
QY  
2702 TAAATCCATGCTCTGGACGCTCCAGAGTGTGGTGTGTTCCAGCTGTTTAGATGAAA 2761  
Db  
2925 TAAATCCATGCTCTGGACGCTCCAGAGTGTGGTGTGTTCCAGCTGTTTAGATGAAA 2984  
QY  
2762 ATCTAAAGATGATATCTTTGAGGAAATCATGGAAGAAATATGAGATATTTAGACAGACC 2821  
Db  
2985 ATCTAAAGATGATATCTTTGAGGAAATCATGGAAGAAATATGAGATATTTAGACAGACC 3044



QY 242 CCAGGCGCTGAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCATTACC 301  
DB 465 CCAGGCGCTGAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCATTACC 524  
QY 302 AAATCCATAAGGAATACCTGCTGGCTGGGCGAGATCATTTGAAACAAATACATTTTAGCA 361  
DB 525 AAATCCATAAGGAATACCTGCTGGCTGGGCGAGATCATTTGAAACAAATACATTTTAGCA 584  
QY 362 GCACTAGTATTTGCCAAGCTGACTATGGCCCTTGAACACTTTGGCCTACCGGATGAACATGT 421  
DB 585 GCACTAGTATTTGCCAAGCTGACTATGGCCCTTGAACACTTTGGCCTACCGGATGAACATGT 644  
QY 422 GCTCTGACAGAGTGGCCAGAAAAGCTGCCAGAGAGTAACTCTCCAGACAGGAATTAAGA 481  
DB 645 GCTCTGACAGAGTGGCCAGAAAAGCTGCCAGAGAGTAACTCTCCAGACAGGAATTAAGA 704  
QY 482 GGTTTGTGGCAGGGGCTCTGGGCTCCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG 541  
DB 705 GGTTTGTGGCAGGGGCTCTGGGCTCCGACTAATAAGACACTCTCTGTGTCCCACTCTGTGG 764  
QY 542 AAAGGCGGATTAAGAAACATCACTTTGATGAGCTTTGTAAGCATACCAAGAGCAGG 601  
DB 765 AAAGGCGGATTAAGAAACATCACTTTGATGAGCTTTGTAAGCATACCAAGAGCAGG 824  
QY 602 CCAAAGGACTCTTGATGGGCGGGTTCATCTTACTCATTTGAAACTATTTTGTACTG 661  
DB 825 CCAAAGGACTCTTGATGGGCGGGTTCATCTTACTCATTTGAAACTATTTTGTACTG 884  
QY 662 CCAATGCCAAGGAGCCTCTTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
DB 885 CCAATGCCAAGGAGCCTCTTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 944  
QY 722 GGCTATCTTTATTTCAAGGACAGATCGTTGATAAAAGTGGCGGACTCTTTCCGGACAGA 781  
DB 945 GGCTATCTTTATTTCAAGGACAGATCGTTGATAAAAGTGGCGGACTCTTTCCGGACAGA 1004  
QY 782 CAGAGAGGATTTGTCATCAGCGTGTCTCATGAGAACCACTCTGTCATTTGATTAAT 841  
DB 1005 CAGAGAGGATTTGTCATCAGCGTGTCTCATGAGAACCACTCTACATTTGGATTAAT 1064  
QY 842 GTCCTTTGGCTGACGTGAGTGAACCTTTTATGAAATTAATTTGAAATGTAACACAG 901  
DB 1065 GTGCTTTGGCTGACGTGAAATGAGACTTTTATTTGAAATTAATTTGAAATGTAACACAG 1124  
QY 902 CCTATGCTCTGTATTCCTCAATGACGCTCTTCCAAACACTTTTGTGACTATGATGAAA 961  
DB 1125 CCTATGCTCTGTATTCCTCAATGACGCTCTTCCAAACACTTTTGTGACTATGATGAAA 1184  
QY 962 CGCCTTCTATGATGGCCAAAGCCTAAAGATTTTGTATGGATGGCTTGGTCAATATAG 1021  
DB 1185 CGCCTTCTATGATGGCCAAAGCCTAAAGATTTTGTATGGATGGCTTGGTCAATATAG 1244  
QY 1022 TTGGAGAGTCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
DB 1245 TTGGAGAGTCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGTGGTC 1141  
DB 1305 ATTGTAAAGCTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGTGGTC 1364  
QY 1142 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAATG 1201  
DB 1365 TAGAGCCCTTCAGGATTTGACCGGTACACCAACTTTTGAACATTTGGAGAGCGCTGTAATG 1424  
QY 1202 TTGCAGATCAGGAAGTTTGTAACTCATATGCGAGGAATATGAAGAGCCCTGTGT 1261  
DB 1425 TTGCAGATCAGGAAGTTTGTAACTCATATGCGAGGAATATGAAGAGCCCTGTGT 1484  
QY 1262 GTGTTGCAAGTGCAGTGAATGGAGCCAGGCTGTTGATGTCACATGATGATG 1321  
DB 1485 GTGTTGCAAGTGCAGTGAATGGAGCCAGGCTGTTGATGTCACATGATGATG 1544

QY 1322 GCATGCTAGATGGTCCAAGTGCATGACAGACATTTTGCACATTAATTTGCTTCGAGCCAG 1381  
DB 1545 GCATGCTAGATGGTCCAAGTGCATGACAGACATTTTGCACATTAATTTGCTTCGAGCCAG 1604  
QY 1382 ACATCCAAAGGTACCTTTGTGCTCATCGACTCTCTCCAAATTTTGTGATTTGAAGCTGGGT 1441  
DB 1605 ACATCCAAAGGTACCTTTGTGCTCATCGACTCTCTCCAAATTTTGTGATTTGAAGCTGGGT 1664  
QY 1442 TAAAGTCTGCCAAGGAAAGTGCAATTTGTCAAATAGCATTTAGTCTGAAAGGAGGAGAGCG 1501  
DB 1665 TAAAGTCTGCCAAGGAAAGTGCAATTTGTCAAATAGCATTTAGTCTGAAAGGAGGAGAGCG 1724  
QY 1502 ACTTCTTTGAGAGGCGCAGGAGATTTAAAGATGATGAGCTGCTATGCTGCTCATGGCTT 1561  
DB 1725 ACTTCTTTGAGAGGCGCAGGAGATTTAAAGATGATGAGCTGCTATGCTGCTCATGGCTT 1784  
QY 1562 TTGATGAAGAGGACAGGCAACAGACACAGACACAAATATCAGAGTGTGCACCCGGGCT 1621  
DB 1785 TTGATGAAGAGGACAGGCAACAGACACAGACACAAATATCAGAGTGTGCACCCGGGCT 1844  
QY 1622 ACCATCTGCTTTGTAAGAAAACCTGGGCTTTAATCCAAATGACATTTATTTTGACCTTAATA 1681  
DB 1845 ACCATCTGCTTTGTAAGAAAACCTGGGCTTTAATCCAAATGACATTTATTTTGACCTTAATA 1904  
QY 1682 TCTTAACCATTTGGGACTGGAATGGAGAACACAACCTTTGTATGCCATTAATTTTATCCATG 1741  
DB 1905 TCTTAACCATTTGGGACTGGAATGGAGAACACAACCTTTGTATGCCATTAATTTTATCCATG 1964  
QY 1742 CAAACAAAGTCATTAAGAAAACATTAACCTGGAGCCAGATTAAGTGAGAGTCTTTTCCAACT 1801  
DB 1965 CAAACAAAGTCATTAAGAAAACATTAACCTGGAGCCAGATTAAGTGAGAGTCTTTTCCAACT 2024  
QY 1802 TGTCTCTCTCTTCCGAGGAATGGAAAGCCATTTGGAAGCAATGCAATGCGGTTTTTCCCTT 1861  
DB 2025 TGTCTCTCTCTTCCGAGGAATGGAAAGCCATTTGGAAGCAATGCAATGCGGTTTTTCCCTT 2084  
QY 1862 ACCATGCAATCAAGTCTGCAATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTGT 1921  
DB 2085 ACCATGCAATCAAGTCTGCAATGGACATGAGATAGTGAATGCTGGAACCTCCCTGTGT 2144  
QY 1922 ATGATGATATCCATAAGGAACCTTTCAGCTCTGTGCAAGATCTCATCTGGAATAAAGACC 1981  
DB 2145 ATGATGATATCCATAAGGAACCTTTCAGCTCTGTGCAAGATCTCATCTGGAATAAAGACC 2204  
QY 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2041  
DB 2205 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2264  
QY 2042 TCATTCAGACTGATGATGGAGAAATGGCCCTGTGCAAGAACGCCCTTGAGTATGCCCTTG 2101  
DB 2265 TCATTCAGACTGATGATGGAGAAATGGCCCTGTGCAAGAACGCCCTTGAGTATGCCCTTG 2324  
QY 2102 TGAAGGGCATTTGAAAACATATTTATTTGAGGATACTGAGGAAGCCAGTGTAAACCAAAAA 2161  
DB 2325 TGAAGGGCATTTGAAAACATATTTATTTGAGGATACTGAGGAGCCAGTGTAAACCAAAAA 2384  
QY 2162 AATATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGGAATGAAATTTGTTG 2221  
DB 2385 AATATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGGAATGAAATTTGTTG 2444  
QY 2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGTGTATTAAGTCAAGCCCGGTTA 2281  
DB 2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGTGTATTAAGTCAAGCCCGGTTA 2504  
QY 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTTCATGGAATAAAGAGAGAGAAACAGAG 2341  
DB 2505 TGAAGAGGCTGTGTGGCCACCTTATCCCTTTTCATGGAATAAAGAGAGAGAAACAGAG 2364  
QY 2342 TGTCTAACCGCACAGTAGAAGAGAGGACCCCTTTACAGGACCACTGCTGCTGCCACTG 2401  
DB 2565 TGTCTAACCGCACAGTAGAAGAGAGGACCCCTTTACAGGACCACTGCTGCTGCCACTG 2624  
QY 2402 TTAAGGCGCAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTTGGCTGCAATA 2461



Db 2625 TTAAGGCGAGCTGCGACGATAGGCGAGAACATAGTTGGAGTAGTCCCTTGGCTGCAATA 2684  
Qy 2462 ATTTCGAGTTATGATTTTAGAGTCAATGACTCATGATGATAGATACTGAAAGCTGCTC 2521  
Db 2685 ATTTCGAGTTATGATTTTAGAGTCAATGACTCATGATGATAGATACTGAAAGCTGCTC 2744  
Qy 2522 TTGACCAACGAGGATATATTTGGCTGTGAGAGTCTCAGACTCATCACTCCCTCCCTCGATGAA 2581  
Db 2745 TTGACCAACGAGGATATATTTGGCTGTGAGAGTCTCAGACTCATCACTCCCTCCCTCGATGAA 2804  
Qy 2582 TGATTTTGTGTCAGAGAAATGAGAGATAGTCTATAAGGATTTCCATGTTGATTTGGAG 2641  
Db 2805 TGATTTTGTGTCAGAGAAATGAGAGATAGTCTATAAGGATTTCCATGTTGATTTGGAG 2864  
Qy 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACTTG 2701  
Db 2865 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACTTG 2924  
Qy 2702 TAATCCATGTCCTGGACCGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761  
Db 2925 TAATCCATGTCCTGGACCGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2984  
Qy 2762 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACC 2821  
Db 2985 ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGATATGAAGATATTAGACAGGACC 3044  
Qy 2822 ATTATGAGTCTCTCAAGGAGGAGATACCTTAAAGCTTAAAGTCAAGCCAGAGAAAGTGTT 2881  
Db 3045 ATTATGAGTCTCTCAAGGAGGAGATACCTTAAAGCTTAAAGTCAAGCCAGAGAAAGTGTT 3104  
Qy 2882 TCCAAATGGATTGGCTGTCTGAACCTCACCGAGTGAAGCCACGTTTATTGGGACCCAGG 2941  
Db 3105 TCCAAATGGATTGGCTGTCTGAACCTCACCGAGTGAAGCCACGTTTATTGGGACCCAGG 3164  
Qy 2942 TCTTTGAAGACTATGACTCGAGAACTGGTGGACTACATTGACTGGAAGCCTTCTTTTG 3001  
Db 3165 TCTTTGAAGACTATGACTCGAGAACTGGTGGACTACATTGACTGGAAGCCTTCTTTTG 3224  
Qy 3002 ATGTCTGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTCCCAAGATATTTAAACACA 3061  
Db 3225 ATGTCTGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTCCCAAGATATTTAAACACA 3284  
Qy 3062 AAACAGTAGTGGAGGCGCAGGAAGTCTACGATGATGCCCAATATGCTGACACAC 3121  
Db 3285 AAACAGTAGTGGAGGCGCAGGAAGTCTACGATGATGCCCAATATGCTGACACAC 3344  
Qy 3122 TGATTAGTCAAAAGAACTCCGGGCCGGGGTGTGGTGGTCTGCGCCAGCACAGAGTA 3181  
Db 3345 TGATTAGTCAAAAGAACTCCGGGCCGGGGTGTGGTGGTCTGCGCCAGCACAGAGTA 3404  
Qy 3182 TCCAGACGACATTCACCTGTACGAGAGGCTGTGTCGCCCGAGCTGCGAGCCCATAG 3241  
Db 3405 TCCAGACGACATTCACCTGTACGAGAGGCTGTGTCGCCCGAGCTGCGAGCCCATAG 3464  
Qy 3242 CCACCTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGAGCCATCT 3301  
Db 3465 CCACCTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGCGAGCCATCT 3524  
Qy 3302 ACTGCCTCTCAGACTTCATCGCTCCCTGTCATCTTGGCATCTGCGATCCGTGCTACCTGGGCTGT 3361  
Db 3525 ACTGCCTCTCAGACTTCATCGCTCCCTGTCATCTTGGCATCTGCGATCCGTGCTACCTGGGCTGT 3584  
Qy 3362 TTGCGGTTGCCCTGCTTTGGGTTAGAGAGCTGAGCAAGGCTTATGAGGATGATGTTGACG 3421  
Db 3585 TTGCGGTTGCCCTGCTTTGGGTTAGAGAGCTGAGCAAGGCTTATGAGGATGATGTTGACG 3644  
Qy 3422 ACTACAGAGCATTCATGTTCAAGGGCTGGGGGACCGGCTGGCAGAGGCCCTTTGACAGAG 3481  
Db 3645 ACTACAGAGCATTCATGTTCAAGGGCTGGGGGACCGGCTGGCAGAGGCCCTTTGACAGAG 3704  
Qy 3482 AGCTCCATGAAGAGTTCCCGCAGAACTGTGGGCTACTGTGGCAGTGTGAGCAGCTGGACG 3541

Db 3705 AGCTCCATGAAGAGTTTCCCGGAGAACTGTGGCCCTACTGTGGCAGTGTGACAGCTGGACG 3764  
Qy 3542 TCGCAGACTCGGAGAGTTTGGGTCAAGGGGATCGCCCGCTCTGCTGCTACCCAGCC 3601  
Db 3765 TCGCAGACTCGGCGAGGCTGCGGTACAGGGGATCGCCCGCTCTGCTGCTACCCAGCC 3824  
Qy 3602 AGCCCGACCAACCCAGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTTACAG 3661  
Db 3825 AGCCCGACCAACCCAGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTTACAG 3884  
Qy 3662 GAATTAGGTTAACAGAAATCATTAGCAATGGCACTGCTTTCAGCAGTCTCAGGCTCTACT 3721  
Db 3885 GCATTAGGTTAACAGAAATCATTAGCAATGGCACTGCTTTCAGCAGTCTCAGGCTCTACT 3944  
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTTCCAAAGGATCAGGTTGAGG 3781  
Db 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTTCCAAAGGATCAGGTTGAGG 4004  
Qy 3782 ATTATGATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
Db 4005 ATTATGATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4064  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTT 3875  
Db 4065 TGGGATATGATACAGACTAACTTTTTTTTTTTT 4098

RESULT 8  
HSU75743 7122 bp mRNA linear PRI 06-JUN-1997  
LOCUS Human methionine synthase mRNA, complete cds.  
DEFINITION  
ACCESSION U75743  
VERSION U75743.1 GI:1763339  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 7122)  
Li, Y.N., Gulati, S., Baker, P.J., Brody, L.C., Banerjee, R. and  
Kruger, W.D.  
Cloning, mapping, and RNA analysis of the human methionine synthase  
gene  
Hum. Mol. Genet. (1996) In press  
REFERENCE 2 (bases 1 to 7122)  
Li, Y.N., Gulati, S., Baker, P.J., Brody, L.C., Banerjee, R. and  
Kruger, W.D.  
Direct Submission  
Submitted (23-OCT-1996) Population Science, Fox Chase Cancer  
Center, 7701 Burholme Ave., Philadelphia, PA 19111, USA  
FEATURES  
source  
1. 7122  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1q42-44"  
/cell\_type="lymphoid"  
/note="composite cDNA"  
287..4084  
/note="similar to Escherichia coli methionine synthase  
encoded by GenBank Accession Number J04975"  
/codon\_start=1  
/product="methionine synthase"  
/protein\_id="AAB58906.1"  
/db\_xref="GI:2160699"  
translation="MSPALDLSQPEGLKKTLDENAIILQKIMVLDDGMGTIOIRE  
KLNEHFRGGEFKDHPRLKNNDDILISITQDVIYQHKYLLAGADIIETNPFSSIS  
IAQDYGLEHLAYEOMCSGAVKAAEEVTLQGIKRFVAGALGPNKTLVSVPVS  
RPYRNTTFDELVEAQAGLLDGGVDILLIETTFDTANAKAALFALQNLPEEKYA  
RPRTFTSGTIVDKSGRTLSQGTGEGFIVSHGEPYIGLNLCALGAEMRPFBIIOK  
CTTAYVLCYPNAGLNTFGDYDETPTSMMAKHLKDFAMDGLVNIIVGGCCGTPDHIREI

CDS



AEAVONCKPRVPATAFEGHMLLSGLEPPRTGPTYNFVNIGERCNVAGSRKPAKLMA  
GNEEALCVAKVOWENMAGQVLVDNMDGMLDPSAMTRFCNLIASBEPDILAKVPLCID  
SNFAVTEAGLKCOQGRKYNLSLXKGEEDDLFEKARKITGKGAAMVMAFDEBQOATE  
TDLKIRVCTVATHLVKGLDGFNPNDIIPDNLTITGKSHMDNIYAINFTHATVKVKE  
TLPGARISGGLSNLSFGPMGEAIREAMHNVPLTHAIKSGMDMGINAGNVLVDDHID  
KELLQICEDLILWNKQDEPAETKLLRYAVAQGTGGKKVITODESWRNGVEERLEVALVKG  
LEKHILIEDTEEARLNKQKPYRPLNLTIEBGLPMANGKICVDLFGAGKMFQVIVISARVM  
IKAVGHILPMEKERESETVNLVVEEDPYOGTIVLATVAGXVDHDKGNIVGVGLG  
NNFRVDLGMVTCVKLAKAALDHKADILGUSGLTSPDSDMIIPVAKEMERLARIPL  
LIGASATSKTHTAPKLAPYASAPVTHLQSGGLSSVVCQSOLDENIEFEETMEIBEY  
DIRQHDYHSLVPLPLSQAQSGFMDLSSPEHPVPTFGITQVDEVDLQKLVGD  
IDMKPFDVWLQKRGYPNGFPFKIFNDKTVGGEARKVYDDAHNNLTLISQKLRADY  
VWGEFPAQSIQDDHILYASAAVPOAABPIATFYGLRQQAQKDSASTEPYICLSDFTAP  
LHSGIRDFGVAFVACFVGEELSKAYEDDGDYSSIMVKALDRLASAPAEABEHVR  
RELWAGYGSQBLDARHLRYKGRIPAPGYSPDHPKTEJTMWRLLADIBQSTGIRLT  
ESLMAAPASAVSGLYFSNLKSKYFAVGKISQOQVEDYALRKNISVAEYEVKWLGFILCY  
DTD"

## ORIGIN

Query Match 89.7%; Score 3517; DB 8; Length 7122;  
Best Local Similarity . 99.8%; Pred. No. 0;  
Matches 3867; Conservative 0; Mismatches 7; Indels 0

Qy	2	GTCACTGTGGAGGACAGCTCTTCTCTGCGCGCCCTCTGCGCAAGAGGAGACTTCGACA	61
Db	225	GTCACTGTGGAGGACAGCTCTTCTCTGCGCGCCCTCTGCGCAAGAGGAGACTTCGACA	284
Qy	62	ACATGTCACCGCGCTCCAAGACTGTCGCAACCGAAGGTCTGAGAAACCCCTGCGGG	121
Db	285	ACATGTCACCGCGCTCCAAGACTGTCGCAACCGAAGGTCTGAGAAACCCCTGCGGG	344
Qy	122	ATGAGATCAATGCCATTCTGCAGAAAGGATATGTGCTGGATGGAGGATGGGACCA	181
Db	345	ATGAGATCAATGCCATTCTGCAGAAAGGATATGTGCTGGATGGAGGATGGGACCA	404
Qy	182	TGATCCAGCGGGAGAAAGCTAAACGAAACAACCTTCGAGGTCAGGAATTTAAAGATCATG	241
Db	405	TGATCCAGCGGGAGAAAGCTAAACGAAACAACCTTCGAGGTCAGGAATTTAAAGATCATG	464
Qy	242	CGAGGCGCTGAAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCAATTACC	301
Db	465	CGAGGCGCTGAAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCAATTACC	524
Qy	302	AAATCCATAGGAATACTTCTGCTGGCGCAGATATCAATTGAACAATACTTTTAAAGCA	361
Db	525	AAATCCATAGGAATACTTCTGCTGGCGCAGATATCAATTGAACAATACTTTTAAAGCA	584
Qy	362	GCAC TAGTATTTGCCCAAGCTGACATATGGCTTTGAACACACTTGGCCCTACCGGATGAACATGT	421
Db	585	GCAC TAGTATTTGCCCAAGCTGACATATGGCTTTGAACACACTTGGCCCTACCGGATGAACATGT	644
Qy	422	GCTCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	481
Db	645	GCTCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA	704
Qy	482	GGTTTGTGGCAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTGCCCATCTGTGG	541
Db	705	GGTTTGTGGCAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTGCCCATCTGTGG	764
Qy	542	AAAGGCCGGAATTATAGGAACATCACATTTGATGAGCTTGTGAGCATATCCNAGAGCAGG	601
Db	765	AAAGGCCGGAATTATAGGAACATCACATTTGATGAGCTTGTGAGCATATCCNAGAGCAGG	824
Qy	602	CCAAGGACCTTCGGATGGGGGTTCGATATCTTTACTCATTTGAACCTATTTTGTGATACTG	661
Db	825	CCAAGGACCTTCGGATGGGGGTTCGATATCTTTACTCATTTGAACCTATTTTGTGATACTG	884
Qy	662	CCAATGCCAAGGCAGCCTTGTGTGCACTCCAAAATCTTTTGTAGGAGAAATATGCTCCCC	721
Db	885	CCAATGCCAAGGCAGCCTTGTGTGCACTCCAAAATCTTTTGTAGGAGAAATATGCTCCCC	944
Qy	722	GGCCTATCTTTTATTTTCAGGGACGATCGTTGATAAAGGTGGCGGACTCTTTTCGGACAGA	781



ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 7224)
TITLE	Yuan, C.-S.
JOURNAL	Methods and compositions for assaying analytes
FEATURES	Patent: US 6376210-A 6 23-APR-2002; General Atomics, San Diego, CA
source	Location/Qualifiers i. 7224 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Query Match	88.4%; Score 3466; DB 6; Length 7224;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 386; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
QY	2 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61
DB	333 GTCACTGTGGAGAGCAGCTTCTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392
QY	62 ACATGTCAACCCGCGCTCCAAAGACCTGTGCAACCCGAAGGTCTGAAGAAACCCCTCGGG 121
DB	393 ACATGTCAACCCGCGCTCCAAAGACCTGTGCAACCCGAAGGTCTGAAGAAACCCCTCGGG 452
QY	122 ATGAGATCAATGCAATCTTCGAGAAGAGGATTAATGCTGTGGATGAGGAGTGGGACCA 181
DB	453 ATGAGATCAATGCAATCTTCGAGAAGAGGATTAATGCTGTGGATGAGGAGTGGGACCA 512
QY	182 TGAATCAGCGGGAGAGCTTAAACGAAGAACACATTCGAGGTCAAGGAATTAAGATCATG 241
DB	513 TGAATCAGCGGGAGAGCTTAAACGAAGAACACATTCGAGGTCAAGGAATTAAGATCATG 572
QY	242 CCAGGCGGCTGAAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCATTTACC 301
DB	573 CCAGGCGGCTGAAAGGCAACAATGACATTTTAAGTATAACTCAGCCTGATGTCATTTACC 632
QY	302 AATCATTAAGGAATCTTCTGGCTGGGCGAGATATCATTTGAAACAAATATCTTTAGCA 361
DB	633 AATCATTAAGGAATCTTCTGGCTGGGCGAGATATCATTTGAAACAAATATCTTTAGCA 692
QY	362 GCACTAGTATTGCCCAAGCTGACTATGGCTTCAACACTTGGCTACCGATCAACATGT 421
DB	693 GCACTAGTATTGCCCAAGCTGACTATGGCTTCAACACTTGGCTACCGATCAACATGT 752
QY	422 GCTCTCAGGAGTGGCGAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 481
DB	753 GCTCTCAGGAGTGGCGAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 812
QY	482 GGTTCGTGGCAGGGCTCTGGTCCGACTTAATAAGACACTCTCTGTGCCCCATCTGTG 541
DB	813 GGTTCGTGGCAGGGCTCTGGTCCGACTTAATAAGACACTCTCTGTGCCCCATCTGTG 872
QY	542 AAGGCGGATTAAGGACATCACAATTTGATGAGCTTGTGAGCATACCAAGACAGG 601
DB	873 AAGGCGGATTAAGGACATCACAATTTGATGAGCTTGTGAGCATACCAAGACAGG 932
QY	602 CCAAGGACTTCTGGATGGCGGGTGTGATCTTACTCTATTGAAACTATTTTGTACTG 661
DB	933 CCAAGGACTTCTGGATGGCGGGTGTGATCTTACTCTATTGAAACTATTTTGTACTG 992
QY	662 CCAATGCCAAGGAGCCTTGTGTCATCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721
DB	993 CCAATGCCAAGGAGCCTTGTGTCATCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052
QY	722 GGCCTATCTTTATTCAGGAGCATCGTTGATTAAGAGTGGCGGACTCTTTTCGGACAGA 781
DB	1053 GGCCTATCTTTATTCAGGAGCATCGTTGATTAAGAGTGGCGGACTCTTTTCGGACAGA 1112
QY	782 CAGGAGGAGATTTGTTCATCAGCGTGTCTCATGAGAACCACTCTCATTTGGATTAAT 841
DB	1113 CAGGAGGAGATTTGTTCATCAGCGTGTCTCATGAGAACCACTCTCATTTGGATTAAT 1172
QY	842 GTGCTTTGGTGCAGCTGAGATGAGACCTTTTATTGAAATTAATTGAAATATGTACACAG 901
DB	1173 GTGCTTTGGTGCAGCTGAAATGAGACCTTTTATTGAAATTAATTGAAATATGTACACAG 1232
QY	902 CTTATGTCCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTGCTGACTATGATGAA 961
DB	1233 CTTATGTCCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTGCTGACTATGATGAA 1292
QY	962 CGCCTTCTATGATGCGCAAGCACTTAAGGATTTTGTCTATGATGCTTGGTCAATATAG 1021
DB	1293 CGCCTTCTATGATGCGCAAGCACTTAAGGATTTTGTCTATGATGCTTGGTCAATATAG 1352
QY	1022 TTGGAGGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTCAAAA 1081
DB	1353 TTGGAGGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTCAAAA 1412
QY	1082 ATTGTAAGCCTTAGAGTTCCACCTGCGACCTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1141
DB	1413 ATTGTAAGCCTTAGAGTTCCACCTGCGACCTGCTTTTGAAGGACATATGTTACTGTCTGGTC 1472
QY	1142 TAGAGCCCTTCAGGATTTGACCCGTACACCACTTTTGTAACTTTGAGAGCGCTGTAATG 1201
DB	1473 TAGAGCCCTTCAGGATTTGACCCGTACACCACTTTTGTAACTTTGAGAGCGCTGTAATG 1532
QY	1202 TTGCAGGATCAAGGAATTTGCTAAACTCATATGCGCAGGAACTATGAAGAGCCTTGT 1261
DB	1533 TTGCAGGATCAAGGAATTTGCTAAACTCATATGCGCAGGAACTATGAAGAGCCTTGT 1592
QY	1262 GTGTTGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGAGATG 1321
DB	1593 GTGTTGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGAGATG 1552
QY	1322 GCATGCTAGATGCTCCAGTGCATGACAGCAATTTTGCAACTTAATTTGCTTCCGAGCCAG 1381
DB	1653 GCATGCTAGATGCTCCAGTGCATGACAGCAATTTTGCAACTTAATTTGCTTCCGAGCCAG 1712
QY	1382 ACATGCGAAGGTACTTTTGTGATGACTCTCTCAATTTTGTCTGCTGATGATGAAGCTGGGT 1441
DB	1713 ACATGCGAAGGTACTTTTGTGATGACTCTCTCAATTTTGTCTGCTGATGATGAAGCTGGGT 1772
QY	1442 TAAAGTGTCCCAAGGAGTGCATTTGCAATAGCATTAGTCTGGAAGGAGGAGAGACG 1501
DB	1773 TAAAGTGTCCCAAGGAGTGCATTTGCAATAGCATTAGTCTGGAAGGAGGAGAGACG 1832
QY	1502 ACTTCTTGGAAGGCGCAGGAAGATTAAAGAGTATGAGCTGCTATGTTGGTTCATGGCTT 1561
DB	1833 ACTTCTTGGAAGGCGCAGGAAGATTAAAGAGTATGAGCTGCTATGTTGGTTCATGGCTT 1892
QY	1562 TTGATGAAGAAGGACAGGCAACAGACACAGACAAAATCAGAGTGTGCAACCCGGGCT 1621
DB	1893 TTGATGAAGAAGGACAGGCAACAGACACAGACAAAATCAGAGTGTGCAACCCGGGCT 1952
QY	1622 ACCATCTGCTGTGAAGAACTGGGCTTTTAAATCCAAATGACATTTTGTGACCTTAATA 1681
DB	1953 ACCATCTGCTGTGAAGAACTGGGCTTTTAAATCCAAATGACATTTTGTGACCTTAATA 2012
QY	1682 TCTTAACCAATTCGGACTTGAATGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 1741
DB	2013 TCTTAACCAATTCGGACTTGAATGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 2072
QY	1742 CAACAAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAGTGAGGTCTTTTCCAACT 1801
DB	2073 CAACAAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAGTGAGGTCTTTTCCAACT 2132
QY	1802 TGTCTCTTCTCTTCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTTCCTTT 1861
DB	2133 TGTCTCTTCTCTTCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGGTTTTCCTTT 2192
QY	1862 ACCATGCAATCAAGTCTGGCATGGAACATGAGATAGTGAATGCTGGAACCTTCCCTGTGT 1921
DB	2193 ACCATGCAATCAAGTCTGGCATGGAACATGAGATAGTGAATGCTGGAACCTTCCCTGTGT 2252
QY	1922 ATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTGGAATAAGAC 1981

Db 2253 ATGATGATATCCATAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTGGAAATAAGACC 2312  
Qy 1982 CTGAGGCCACTGAGAGAGCTCTTACGTTATGCCCCAGACTCAAGGCAACAGAGGAGGAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAGAGCTCTTACGTTATGCCCCAGACTCAAGGCAACAGAGGAGGAGAAAG 2372  
Qy 2042 TCATTGAGCTGATGAGTGAAGAAATGGCCCTCTCGAAGAACCCCTTTGAGTATGCCCTTG 2101  
Db 2373 TCATTGAGCTGATGAGTGAAGAAATGGCCCTCTCGAAGAACCCCTTTGAGTATGCCCTTG 2432  
Qy 2102 TGAAGGGCAATTGAAAAACATATTATTGAGGATATCTGAGGAAGCCAGGTTAAACCAAAAAA 2161  
Db 2433 TGAAGGGCAATTGAAAAACATATTATTGAGGATATCTGAGGAAGCCAGGTTAAACCAAAAAA 2492  
Qy 2162 AATATCCCGACCTCTCAATATATAATGAAGGACCCCTGATGGAATGGAATGGAATTTGTTG 2221  
Db 2493 AATATCCCGACCTCTCAATATATAATGAAGGACCCCTGATGGAATGGAATGGAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAGAGTCAGCCCGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACTCAGGTTATAGAGTCAGCCCGGTTA 2612  
Qy 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTTCAATGAAAAAGAAAGAGAAAGAACCCAGAG 2341  
Db 2613 TGAAGAGGCTGTGGCCACCTTATCCCTTTTCAATGAAAAAGAAAGAGAAAGAACCCAGAG 2672  
Qy 2342 TGCTTAACGGCACAGTGAAGAGAGGACCCCTTACCAAGGCAACCATCGTCTGCCCACTG 2401  
Db 2673 TGCTTAACGGCACAGTGAAGAGAGGACCCCTTACCAAGGCAACCATCGTCTGCCCACTG 2732  
Qy 2402 TTAAGCGGACGTGCACACATAGGCAAGAACATAGTTGAGTAGTCTTGGCTGCATA 2461  
Db 2733 TTAAGCGGACGTGCACACATAGGCAAGAACATAGTTGAGTAGTCTTGGCTGCATA 2792  
Qy 2462 ATTTCCGAGTTATTGATTTAGGAGTCATGACTCATGTGATAGATATAGATCTGAAGAGCTGCTC 2521  
Db 2793 ATTTCCGAGTTATTGATTTAGGAGTCATGACTCATGTGATAGATATAGATCTGAAGAGCTGCTC 2852  
Qy 2522 TTGACCAAAAGCAGATATATAATGGCCCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2581  
Db 2853 TTGACCAAAAGCAGATATATAATGGCCCTGTGAGGACTCATCATCTCTTCCCTGGATGAAA 2912  
Qy 2582 TGATTTTGTGCAAGGAAATGAGAGATTAGCTATAGGATTCATTTGATTTGGAG 2641  
Db 2913 TGATTTTGTGCAAGGAAATGAGAGATTAGCTATAGGATTCATTTGATTTGGAG 2972  
Qy 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTTCCGAGATACAGTGCACTG 2701  
Db 2973 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTTCCGAGATACAGTGCACTG 3032  
Qy 2702 TAATCCATGTCCTGGACGCGTCCAAGAGTGGTGGTGTGTTCCAGCTGTTAGATGAAA 2761  
Db 3033 TAATCCATGTCCTGGACGCGTCCAAGAGTGGTGGTGTGTTCCAGCTGTTAGATGAAA 3092  
Qy 2762 ATCTAAGGATGAATATCTTGAAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGACC 2821  
Db 3093 ATCTAAGGATGAATATCTTGAAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGACC 3152  
Qy 2822 ATATGAGTCTCTCAAGGAGAGGAGATACTTACCTTAAGTCAAGCCAGAAAAAGTGGTT 2881  
Db 3153 ATATGAGTCTCTCAAGGAGAGGAGATACTTACCTTAAGTCAAGCCAGAAAAAGTGGTT 3212  
Qy 2882 TCCAAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGCTTTATTGGGACCCAGG 2941  
Db 3213 TCCAAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCCAAGCTTTATTGGGACCCAGG 3272  
Qy 2942 TCTTTGAAGACTATGACCTGCAGAGAGCTGGTGGACTACATTTGACTGGAGCCTTCTTTG 3001  
Db 3273 TCTTTGAAGACTATGACCTGCAGAGAGCTGGTGGACTACATTTGACTGGAGCCTTCTTTG 3332  
Qy 3002 ATGCTGGCAGCTCCGGGGCAAGTACCCGAATCCGAGGCTTCCCAAGATATTTAACGACA 3061

Db 3333 ATGCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTTCCCAAGATATTTAAACGACA 3392  
Qy 3062 AAACAGTAGTGTGAGAGGCGCAGAAAGGTTCTAGATGATGCCCCACAAATATGCTGAAACACAC 3121  
Db 3393 AAACAGTAGTGTGAGAGGCGCAGAAAGGTTCTAGATGATGCCCCACAAATATGCTGAAACACAC 3452  
Qy 3122 TGATTAGTCAAAAGAAAACTCCGGGCGCGGGGTGTGTGGTTCCTGGCCAGACAGAGTA 3181  
Db 3453 TGATTAGTCAAAAGAAAACTCCGGGCGCGGGGTGTGTGGTTCCTGGCCAGACAGAGTA 3512  
Qy 3182 TCAAGACAGCAATTCACCTGTAGCAGAGGCTGTGTGCCCCCAGGCTGAGAGCCCATAG 3241  
Db 3513 TCAAGACAGCAATTCACCTGTAGCAGAGGCTGTGTGCCCCCAGGCTGAGAGCCCATAG 3572  
Qy 3242 CCACCTTTCTATGGGTTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATACT 3301  
Db 3573 CCACCTTTCTATGGGTTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATACT 3632  
Qy 3302 ACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCTGGCAATCCGTGACTACCTGGGCTGT 3361  
Db 3633 ACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCTGGCAATCCGTGACTACCTGGGCTGT 3692  
Qy 3362 TTGCCGTTGCTGCTTTGGGGTAGAAGACTGAGCAAGGCTTATGAGGATGATGGTAGC 3421  
Db 3693 TTGCCGTTGCTGCTTTGGGGTAGAAGACTGAGCAAGGCTTATGAGGATGATGGTAGC 3752  
Qy 3422 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAGAG 3481  
Db 3753 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAGAG 3812  
Qy 3482 AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGGACG 3541  
Db 3813 AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGGACG 3872  
Qy 3542 TCCAGACCTGCAAGAGTTCGGGTCAAGGCGCATCCGCCGCTCTCTGCTACCCAGCC 3601  
Db 3873 TCCAGACCTGCAAGAGTTCGGGTCAAGGCGCATCCGCCGCTCTCTGCTACCCAGCC 3932  
Qy 3602 AGCCGACACACACCGAGAGGCTCACCATGTGGAGACTCCGACAGATCGAGCAGCTACAG 3661  
Db 3933 AGCCGACACACACCGAGAGGCTCACCATGTGGAGACTTCGACAGATCGAGCAGCTACAG 3992  
Qy 3662 GCATTAGGTTAAACAGAAATCAATTAGCAATGGCACTCTGCTTCCAGCAGTCTCAGGCTCTACT 3721  
Db 3993 GCATTAGGTTAAACAGAAATCAATTAGCAATGGCACTCTGCTTCCAGCAGTCTCAGGCTCTACT 4052  
Qy 3722 TCTCCAAATTTGAAAGTCCAAATATTTTGTGTGGGGAGATTTCCAAAGGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGAAAGTCCAAATATTTTGTGTGGGGAGATTTCCAAAGGATCAGGTTGAGG 4112  
Qy 3782 ATTATGCAATTTGAGGAGAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 3841  
Db 4113 ATTATGCAATTTGAGGAGAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 4172  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 3875  
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 4206

## RESULT 10

AX069342

LOCUS

Sequence 6 from Patent WO0102600.

DEFINITION

AX069342

ACCESSION

AX069342.1

VERSION

GI:12579205

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Suarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1

AUTHORS

Yuan, C. S.





1562 TTGATGAAGAGGACGCGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCT 1621  
1563 TTTGATGAAGAGGACGCGCAACAGAAAACAGACACAAAAATCAGAGTGTGCACCCGGCCCT 1562  
1622 ACCATCTGCTGTGAAAAAACTGGGCTTTTAAATCCAAATGACATTAATTTTGGACCCCTAATA 1681  
1623 ACCATCTGCTGTGAAAAAACTGGGCTTTTAAATCCAAATGACATTAATTTTGGACCCCTAATA 2012  
1682 TCCTAACCATTTGGGACCTGGAAATGGAGGAAACAACTTTGTATGCCATTAATTTTATCCATG 1741  
2013 TCCTAACCATTTGGGACCTGGAAATGGAGGAAACAACTTTGTATGCCATTAATTTTATCCATG 2072  
1742 CAACAAAGATCATTTAAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGCTCTTTCCAACT 1801  
2073 CAACAAAGATCATTTAAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGCTCTTTCCAACT 2132  
1802 TGTCTCTTCTCTTCGAGGAAATGGAAGCAATTCGAGAAAGCAATGCAATGGGGTTTTCCTTT 1861  
2133 TGTCTCTTCTCTTCGAGGAAATGGAAGCAATTCGAGAAAGCAATGCAATGGGGTTTTCCTTT 2192  
1862 ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT 1921  
2193 ACCATGCAATCAAGTCTGGCATGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT 2252  
1922 ATGATGATATCCATAAGGAACCTCTGCAGCTCTCTGGAAGATCTCATCTGGAATAAGAAC 1981  
2253 ATGATGATATCCATAAGGAACCTCTGCAGCTCTCTGGAAGATCTCATCTGGAATAAGAAC 2312  
1982 CTGAGGCCACTGAGAAGCTCTTACGTTATGCCCCAGACTCAAGGCCACAGGAGGGAAGAAAG 2041  
2313 CTGAGGCCACTGAGAAGCTCTTACGTTATGCCCCAGACTCAAGGCCACAGGAGGGAAGAAAG 2372  
2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAATGCTGATGCTCCCTTG 2101  
2373 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGACGCTTGAATGCTGATGCTCCCTTG 2432  
2102 TGAAGGGCATTTGAAAAACATATTTAGAGGATCTGAGGAGCCAGGTTAAACCAAAAAA 2161  
2433 TGAAGGGCATTTGAAAAACATATTTAGAGGATCTGAGGAGCCAGGTTAAACCAAAAAA 2492  
2162 AATATCCCGACCTCTCAATATAATGGAAGACCCCTGATGAATGGAATGGAATGGAATGTTG 2221  
2493 AATATCCCGACCTCTCAATATAATGGAAGACCCCTGATGAATGGAATGGAATGGAATGTTG 2552  
2222 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATTAAGTCAAGCCCGGTTA 2281  
2553 GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATTAAGTCAAGCCCGGTTA 2612  
2282 TGAAGAGGGCTGTGGCCACCTTATCCCTTTTATGGAAGAAAGAGAGAAACCAAGAG 2341  
2613 TGAAGAGGGCTGTGGCCACCTTATCCCTTTTATGGAAGAAAGAGAGAAACCAAGAG 2672  
2342 TGCTTAAACCGGACAGTGAAGAGAGGACCCCTTACAGGGGACCATCGTGTGCGCCACTG 2401  
2673 TGCTTAAACCGGACAGTGAAGAGAGGACCCCTTACAGGGGACCATCGTGTGCGCCACTG 2732  
2402 TTAAGGGCAGCTGCAACAGATAGGCAAGACATAGTGGAGTAGTCTTGGCTGCAATA 2461  
2733 TTAAGGGCAGCTGCAACAGATAGGCAAGACATAGTGGAGTAGTCTTGGCTGCAATA 2792  
2462 ATTTCCGAGTTATTTAGAGTCAAGCTCCATGTGATAAGATCTGAAAGCTGCTC 2521  
2793 ATTTCCGAGTTATTTAGAGTCAAGCTCCATGTGATAAGATCTGAAAGCTGCTC 2852  
2522 TTGACCAAAAGCAGATATAATGGCTGTGAGGACTCATCACTCTTCCCTGATGATAA 2581  
2853 TTGACCAAAAGCAGATATAATGGCTGTGAGGACTCATCACTCTTCCCTGATGATAA 2912  
2582 TGAATTTTGTGTCAGGAAATGGAGATAGCTATAAGGATTTCCATTTGTTGATGGAG 2641  
2913 TGAATTTTGTGTCAGGAAATGGAGATAGCTATAAGGATTTCCATTTGTTGATGGAG 2972  
2642 GAGCAACCACTTTCAAAAAACCCACACAGCAGGTTAAAAATAGCTCCGAGATACAGTGCACCTG 2701

2973 GAGCAACCACTTTCAAAAAACCCACACAGCAGTGTAAAAATAGCTCCGAGATACAGTGCACCTG 3032  
2702 TAATCCATGTCTCGAGCGGTCCAGAGAGTGTGGTGTGTGTTCCAGCTGTTAGATGAAA 2761  
3033 TAATCCATGTCTCGAGCGGTCCAGAGAGTGTGGTGTGTGTTCCAGCTGTTAGATGAAA 3092  
2762 ATCTAAGGATGAATACTTTTGAGGAAATCATGGAAGAAATATGAAGATATTTAGACAGACC 2821  
3093 ATCTAAGGATGAATACTTTTGAGGAAATCATGGAAGAAATATGAAGATATTTAGACAGACC 3152  
2822 ATTATGAGTCTCTCAAGGAGAGGAGATACCTTAACTTAACTCAAGCCAGAAAAAGTGGTT 2881  
3153 ATTATGAGTCTCTCAAGGAGAGGAGATACCTTAACTTAACTCAAGCCAGAAAAAGTGGTT 3212  
2882 TCCAAATGGAATGGCTGTCTGAACCTCAACCCAGTGAAGCCAGTTTATTTGGAGCCAGG 2941  
3213 TCCAAATGGAATGGCTGTCTGAACCTCAACCCAGTGAAGCCAGTTTATTTGGAGCCAGG 3272  
2942 TCTTTGAAGACTATGACTGCGAAGCTGGTGGACTTACATTTGACTGGAAGCCCTTTCTTTG 3001  
3273 TCTTTGAAGACTATGACTGCGAAGCTGGTGGACTTACATTTGACTGGAAGCCCTTTCTTTG 3332  
3002 ATGTCTGGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTAAACGACA 3061  
3333 ATGTCTGGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTAAACGACA 3392  
3062 AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCCCCAATATGCTGAACACAC 3121  
3393 AAACAGTAGTGGAGAGGCGCAGGAAGTCTACGATGATGCCCCCAATATGCTGAACACAC 3452  
3122 TGATTTAGTCAAAAAAATCTCCGGGCGGGGTGGTGGGTTCTGGCCAGACACAGATA 3181  
3453 TGATTTAGTCAAAAAAATCTCCGGGCGGGGTGGTGGGTTCTGGCCAGACACAGATA 3512  
3182 TCCAAGAAGCAGATTTACCTCTGTACGAGAGGCTGCTGTGCCCCCAGGCTGCAAGGCCATAG 3241  
3513 TCCAAGAAGCAGATTTACCTCTGTACGAGAGGCTGCTGTGCCCCCAGGCTGCAAGGCCATAG 3572  
3242 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGAGGACTTGCAGCAGCGAGCCATACT 3301  
3573 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAGAGGACTTGCAGCAGCGAGCCATACT 3632  
3302 ACTGCGCTCTCAGACTTCACTCGCTCCCTTGCATTTCTGGCATCTCGTACTACTCGGCGCTGT 3361  
3633 ACTGCGCTCTCAGACTTCACTCGCTCCCTTGCATTTCTGGCATCTCGTACTACTCGGCGCTGT 3692  
3362 TTGCGGTTGCTCTTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGAACG 3421  
3693 TTGCGGTTGCTCTTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGAACG 3752  
3422 ACTACAGCAGCATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAAGAG 3481  
3753 ACTACAGCAGCATCATGGTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAAGAG 3812  
3482 AGCTTCAATGAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGACAG 3541  
3813 AGCTTCAATGAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGACAG 3872  
3542 TCGCAGACCTGCGAAGGTTTCGGGTACAAAGGCAATCCGCCCGGCTCTGGCTTACCCAGGCC 3601  
3873 TCGCAGACCTGCGAAGGTTTCGGGTACAAAGGCAATCCGCCCGGCTCTGGCTTACCCAGGCC 3932  
3602 AGCCCGACACACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3661  
3933 AGCCCGACACACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3992  
3662 GCATTAGGTTAAACAGAAATCATTTAGCAATGGCAGCTGCTTCAAGAGTCTCAGGCGCTCTACT 3721  
3993 GCATTAGGTTAAACAGAAATCATTTAGCAATGGCAGCTGCTTCAAGAGTCTCAGGCGCTCTACT 4052  
3722 TCTCCAAATTTGAAAGTCCAAATAATTTTGGCTGTGGGGAAAGATTTTCAAGGATCAGGTTGAGG 3781



```
Db 4053 TCCTCAATTTGAAGTCCAAATATTTTCTGCTGGGGGAAGATTTCCAGGATCAGGTTGAGG 4112
Qy 3782 ATTATGCAATTCAGGAAGACATATCTGCTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 3841
Db 4113 ATTATGCAATTCAGGAAGACATATCTGCTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 4172
Qy 3842 TGGATATGATACAGACTAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3875
Db 4173 TGGATATGATACAGACTAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4206

RESULT 11
HSU73338
LOCUS HSU73338 7224 bp mRNA linear PRI 02-JAN-1997
DEFINITION Human methionine synthase mRNA, complete cds.
ACCESSION U73338
VERSION U73338.1 GI:1763268
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 7224)
AUTHORS Chen,L.H., Liu,M.-L., Hwang,H.-Y., Chen,L.-S., Korenberg,J. and
Shane,B.
TITLE Human methionine synthase: cDNA cloning, gene localization and
expression
JOURNAL J. Biol. Chem. (1996) In press
REFERENCE 2 (bases 1 to 7224)
AUTHORS Chen,L.H., Liu,M.-L., Hwang,H.-Y., Chen,L.-S., Korenberg,J. and
Shane,B.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-1996) Nutritional Sciences, University of
California, 117 Morgan Hall, Berkeley, CA 94720-3104, USA
FEATURES
source
location/Qualifiers
1..7224
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1q42.3-1q43"
/tissue_type="liver"
/dev_stage="fetus"
395..4192
/EC_number="2.1.1.13"
/codon_start=1
/product="methionine synthase"
/protein_id="AAB39704.1"
/db_xref="GI:1763269"
/translations="MSPALQDLSPQGLKKLTLDRLINAILQKRMVLDGGMGTMIQRE
KLANEHPRGQEFKHARPLKGNNDILSIITOPDVIYOHKEVLLAGADIIETNTPFSST
IAQADYGLHAYRMWCSAGVAKAAEVTLTGIGKRFVAGALGPNKTLVSPSVE
RPDYNITFDLVEAQEGKGLDGVILLIETITFDYANAKAALFALONLPEEKIA
PRPFIISGTVLKSGLTSQGTGEPFVSVSHGEPICILGNLALGAEMRPFETIEIK
CTTAYLVCPNAGLPNTFGDYDETPSMMAKHLKDFAMDGLVNVGCGCSSTPDHIREI
AEAVKCPKPRPATAPNFGRIKGLLEPPFRIGPYTNFVNIERNVAGSRKPAKLIMA
GNYEALCVAKVQVMEGAOVLVDNMDMDLGDPSAMTRPCNLIASPDIAPLCITDS
SNFAVEAGLKCKQKCIVNSISLKEGEDDFLEKARKIKKYGAAMVMAFDEGQATE
TDTKRVCTRAYLHLLVKGLGFNPDIIIFDNILITGTGMEHNLYAINFTHATVLIK
TLPGARISGLNLSFSFRGMEAIRAMHGFVLYHAIKSGMDMGIYNAGNLFPYDDIH
KELLQCEDLIWNKDEATEKLLRYAQTOGTGGKVTQTDWWRNGPVEERLEYLVKG
IEKHI1EDTEAEALNKKYPRPLNI1EGPLANGMKIVGDLFGAGKMLPQVIKSAVM
KVAVHLI1PWEKERETRVLNGTVREEDPYQCTIVLATVKGDPVHDIGKNI1GVVLGC
NFRVLDLWTHPTDKILKAA1LDHADI1GLSGLI1TPSLDEMI1FVAKEMERLAIR1PL
LI1GGAT1GKMT1AK1P1RY1AP1V1H1VD1ASK1SVV1CS1QLL1DEN1K1DE1FPE1E1ME1SE1Y
DIRQDHYESLKER1Y1PL1SQARKS1GFQ1MDL1SE1P1K1V1P1F1G1Q1V1DY1D1Q1LV1DY
IDWKP1FDM1V1K1P1R1G1P1K1F1ND1K1T1G1V1G1E1A1R1Y1DD1AH1N1M1L1S1Q1K1L1R1G
V1G1F1P1A1Q1S1Q1D1H1L1Y1A1E1A1V1P1O1A1E1P1A1T1F1Y1L1R1Q1A1E1K1S1T1E1P1Y1C1S1D1F1A1P
LH1S1G1D1Y1L1G1F1A1V1A1C1F1G1E1L1S1K1A1V1E1D1D1G1D1Y1S1I1M1K1A1L1D1R1L1A1E1A1P1E1L1H1E1R1V
REL1W1A1Y1C1G1S1Q1D1V1A1D1R1L1Y1K1G1R1P1A1P1Q1P1S1Q1D1H1E1K1I1W1M1L1A1E1Q1S1T1G1R1T
ES1L1A1M1A1P1A1S1A1V1S1G1Y1F1S1N1L1K1S1Y1F1A1V1K1S1Q1D1Y1A1L1R1K1N1I1S1V1A1E1V1E1K1W1G1P1I1L1Y
D1D1"
```

Db	1113	CAGGAGGAGGATTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATGGGATTAAATT	1171
Qy	842	GTGCTTTGGGTGCAGCTGAGATGAGACCTTTTATTGAAATAATGGAAAATGTACAACAG	901
Db	1173	GTGCTTTGGGTGCAGCTGAAATGAGACCTTTTATTGAAATAATGGAAAATGTACAACAG	1232
Qy	902	CCTATGTCCTCTGTTATCCAAATGCAGGTCTTCCCAACACCTTTTGTGTGACTATGATGAAA	961
Db	1233	CCTATGTCCTCTGTTATCCCAATGCAGGTCTTCCCAACACCTTTTGTGTGACTATGATGAAA	1292
Qy	962	CGCCTTCTATGATGGCCCAAGCACCTTAAGGATTTTGTCTATGGATGCGTTGGTCAATATAG	1021
Db	1293	CGCCTTCTATGATGGCCCAAGCACCTTAAGGATTTTGTCTATGGATGCGTTGGTCAATATAG	1352
Qy	1022	TTGAGAGATGCTGTGGGTCAACACCAAGATCATATCAGGAAAATGTCTGAAGCTGTGAAA	1081
Db	1353	TTGAGAGATGCTGTGGGTCAACACCAAGATCATATCAGGAAAATGTCTGAAGCTGTGAAA	1412
Qy	1082	ATTGTAAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGGTC	1141
Db	1413	ATTGTAAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGGTC	1472
Qy	1142	TAGAGCCCTCAGGATTTGGACCGTACACCAACTTTGTTAACATTTGAGAGCGCTGTAATG	1201
Db	1473	TAGAGCCCTCAGGATTTGGACCGTACACCAACTTTGTTAACATTTGAGAGCGCTGTAATG	1532
Qy	1202	TTGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGCAGGAACTATGAAGAAGCCCTTGT	1261
Db	1533	TTGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGCAGGAACTATGAAGAAGCCCTTGT	1592
Qy	1262	GTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCCAAGTGTGGATGTCAACATGGATGATG	1321
Db	1593	GTGTTGCCAAAGTGCAGGTGGAAAATGGGAGCCCAAGTGTGGATGTCAACATGGATGATG	1652
Qy	1322	GCATGCTAGATGGTCCAAGTGCAAATGACAGATTTTGTGCAACTTAATTTGCTTCGAGCCAG	1381
Db	1653	GCATGCTAGATGGTCCAAGTGCAAATGACAGATTTTGTGCAACTTAATTTGCTTCGAGCCAG	1712
Qy	1382	ACATCGCAAAAGGTACCTTTGTGCAATGCATCCCTCCAATTTTGTCTGATGTAAGACTGGT	1441
Db	1713	ACATCGCAAAAGGTACCTTTGTGCAATGCATCCCTCCAATTTTGTCTGATGTAAGACTGGT	1772
Qy	1442	TAAAGTGTCTGCCAAGGGAAGTGCAATTGTCAATAGCATTTAGTCTGGAAGGAAGGAGGACG	1501
Db	1773	TAAAGTGTCTGCCAAGGGAAGTGCAATTGTCAATAGCATTTAGTCTGGAAGGAAGGAGGACG	1832
Qy	1502	ACTTCTTGGAGAAAGCCGAGGAAGATTAAAGATATGGAGCTGCTATGGTGGTCTATGCGCTT	1561
Db	1833	ACTTCTTGGAGAAAGCCGAGGAAGATTAAAGATATGGAGCTGCTATGGTGGTCTATGCGCTT	1892
Qy	1562	TTGATGAGAGGACAGCCACAGAAACAGACACAAAATCAGAGTGTGCACCCGGGCGCT	1621
Db	1893	TTGATGAGAGGACAGCCACAGAAACAGACACAAAATCAGAGTGTGCACCCGGGCGCT	1952
Qy	1622	ACCATCTGCTTGTGAAAATACTGGGCTTTAATCCAAATGACATTAATTTTGTGACCCCTAATA	1681
Db	1953	ACCATCTGCTTGTGAAAATACTGGGCTTTAATCCAAATGACATTAATTTTGTGACCCCTAATA	2012
Qy	1682	TCCTAACCAATTTGGGATGGGAACTGGAAGCAATTCGAGAGCAATTCGATGGGCTTTTCCCTT	1741
Db	2013	TCCTAACCAATTTGGGATGGGAACTGGAAGCAATTCGAGAGCAATTCGATGGGCTTTTCCCTT	2072
Qy	1742	CAACAAAAGTCAATTAAGAAAATTAATCTTGAGCCAGATTAAGTGGAGGTCTTTTCCAACCT	1801
Db	2073	CAACAAAAGTCAATTAAGAAAATTAATCTTGAGCCAGATTAAGTGGAGGTCTTTTCCAACCT	2132
Qy	1802	TGTCCTCTCTCTTCGAGGAATGGAAGCAATTCGAGAGCAATTCGATGGGCTTTTCCCTT	1861
Db	2133	TGTCCTCTCTCTTCGAGGAATGGAAGCAATTCGAGAGCAATTCGATGGGCTTTTCCCTT	2192
Qy	1862	ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGAATGTCTGGAACCTCCCTGTGT	1921
Db	2193	ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGAATGTCTGGAACCTCCCTGTGT	2252

Qy	1922	ATGATGATATCCATAAGAACTTCTCTGAGCTCTGTGAGAGATCTCATCTGGAATAAAAGACC	1981
Db	2253	ATGATGATATCCATAAGAACTTCTCTGAGCTCTGTGAGAGATCTCATCTGGAATAAAAGACC	2312
Qy	1982	CTGAGGCCACTGAGAGAGCTCTTACGTTATGCGCCAGACTCAAGGCA CAGGAGGGAAGAAAG	2041
Db	2313	CTGAGGCCACTGAGAGAGCTCTTACGTTATGCGCCAGACTCAAGGCA CAGGAGGGAAGAAAG	2372
Qy	2042	TCATTCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCCTTG	2101
Db	2373	TCATTCAGACTGATGATGGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCCTTG	2432
Qy	2102	TGAAGGCCATTGAAAACATATTTATGAGGATACTGAGGAGCCAGGTTAAACCAAAAA	2161
Db	2433	TGAAGGCCATTGAAAACATATTTATGAGGATACTGAGGAGCCAGGTTAAACCAAAAA	2492
Qy	2162	AATATCCCAGACTCTCAATAATTAATGAAGGACCCCTGATGAATGGAAATGAAAATGTTTG	2221
Db	2493	AATATCCCAGACTCTCAATAATTAATGAAGGACCCCTGATGAATGGAAATGAAAATGTTTG	2552
Qy	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTA CTTCAAGGTTATAAAGTCHAGCCCGGGTTA	2281
Db	2553	GTGATCTTTTGGAGCTGGAAAAATGTTTCTA CTTCAAGGTTATAAAGTCHAGCCCGGGTTA	2612
Qy	2282	TGAAGAAGGCTGTGGCCACTTATCCCTTTTCAATGGAAGAAAGAGAGAAACACAGAG	2341
Db	2613	TGAAGAAGGCTGTGGCCACTTATCCCTTTTCAATGGAAGAAAGAGAGAAACACAGAG	2672
Qy	2342	TGCTTAACGGCACAGTAGAAGAGAGGACCTTTA CCAAGGACCATCTGCTGGCCACTG	2401
Db	2673	TGCTTAACGGCACAGTAGAAGAGAGGACCTTTA CCAAGGACCATCTGCTGGCCACTG	2732
Qy	2402	TTAAAGCGGAGTGCACGACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGGCAATA	2461
Db	2733	TTAAAGCGGAGTGCACGACATAGGCGAAGAACATAGTTGGAGTAGTCTTGGCTGGCAATA	2792
Qy	2462	ATTTCCGAGTTATTTGATTTAGGAGTCAATGACTCCATGTGATAAGATACTGAAAGCTGCTC	2521
Db	2793	ATTTCCGAGTTATTTGATTTAGGAGTCAATGACTCCATGTGATAAGATACTGAAAGCTGCTC	2852
Qy	2522	TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTGATAGAGATACTGAAAGTGA	2581
Db	2853	TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTGATAGAGATACTGAAAGTGA	2912
Qy	2582	TGATTTTGTGGCCAGGAAATGGAGAGATTAGCTATAAGGATTTCCATTGCTGATTTGGAG	2641
Db	2913	TGATTTTGTGGCCAGGAAATGGAGAGATTAGCTATAAGGATTTCCATTGCTGATTTGGAG	2972
Qy	2642	GAGCAACCACTTTCAAAAACCCACACAGCAGTTAAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2973	GAGCAACCACTTTCAAAAACCCACACAGCAGTTAAAATAGCTCCGAGATACAGTGCACCTG	3032
Qy	2702	TAATCCATGCTCGACGCTCAGAGAGTGTGGTGTGTTCCAGCTCTTATAGATGAAA	2761
Db	3033	TAATCCATGCTCTCGACGCTCAGAGAGTGTGGTGTGTTCCAGCTCTTATAGATGAAA	3092
Qy	2762	ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATGAAGATATGAAGATATTAGACGAC	2821
Db	3093	ATCTAAAGGATGAATACCTTTGAGGAAATCATGGAAGAAATGAAGATATTAGACGACGAC	3152
Qy	2822	ATTATGAGTCTCTCAAGGAGAGGAGATATCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT	2881
Db	3153	ATTATGAGTCTCTCAAGGAGAGGAGATATCTTACCTTTAAGTCAAGCCAGAAAAAGTGGTT	3212
Qy	2882	TCCAAATCGATTTGGCTGTCTCAACCTCACCCAGTGAAGCCACCTTTATTTGGGACCCAGG	2941
Db	3213	TCCAAATCGATTTGGCTGTCTCAACCTCACCCAGTGAAGCCACCTTTATTTGGGACCCAGG	3272
Qy	2942	TCTTTGAAGACTATGACTGAGAGCTGGTGGACTCAATTTGACTGGAAGCTTTCTTTTG	3001
Db	3273	TCTTTGAAGACTATGACTGAGAGCTGGTGGACTCAATTTGACTGGAAGCTTTCTTTTG	3332

Qy	3002	ATGTCGGCAGCTCCGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTAAGACA	3061
Db	3333	ATGTCGGCAGCTCCGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTAAGACA	3392
Qy	3062	AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCACAATATGCTGAACACAC	3121
Db	3393	AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCACAATATGCTGAACACAC	3452
Qy	3122	TGATTAGTCAAAAGAACTCCGGGCCGGGGTGTGGTGGTCTGGCCAGCAGAGTA	3181
Db	3453	TGATTAGTCAAAAGAACTCCGGGCCGGGGTGTGGTGGTCTGGCCAGCAGAGTA	3512
Qy	3182	TCCAAGACGACATTCACCTGTACGAGAGCTGTGTGCCCGCCAGGCTGCGAGGCCATAG	3241
Db	3513	TCCAAGACGACATTCACCTGTACGAGAGCTGTGTGCCCGCCAGGCTGCGAGGCCATAG	3572
Qy	3242	CCACTTTCTATGGGTTAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATCT	3301
Db	3573	CCACTTTCTATGGGTTAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATCT	3632
Qy	3302	ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCTGGCATCCGTGACTACCTGGGCTGT	3361
Db	3633	ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCTGGCATCCGTGACTACCTGGGCTGT	3692
Qy	3362	TTGCCGTGCTCTTGGGGTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG	3421
Db	3693	TTGCCGTGCTCTTGGGGTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG	3752
Qy	3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTCGAGAG	3481
Db	3753	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCTTTCGAGAG	3812
Qy	3482	AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTACTGTGCGAGTGAGCAGCTGGAGC	3541
Db	3813	AGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTACTGTGCGAGTGAGCAGCTGGAGC	3872
Qy	3542	TCCAGACCTCGAAGGTTGGGCTACAGGGCATCCGCGGCTCTGGGTACCCAGCC	3601
Db	3873	TCCAGACCTCGCAGGCTCGGCTACAGGGGCATCCGCGGCTCTGGGTACCCAGCC	3932
Qy	3602	AGCCCGACACACCGAAGCTCACATGTGGAGCTTCGAGACATCGCAGAGCTCTACAG	3661
Db	3933	AGCCCGACACACCGAAGCTCACATGTGGAGCTTCGAGACATCGCAGAGCTCTACAG	3992
Qy	3662	GCATTAGGTTAAGCAATATTAGCAATGGCACCTGTCTCAGCAGTCTCAGGCTCTACT	3721
Db	3993	GCATTAGGTTAAGCAATATTAGCAATGGCACCTGTCTCAGCAGTCTCAGGCTCTACT	4052
Qy	3722	TCCTCAATTTGAAGTCCAAATATTTGCTGTGGGAAGATTTCCAGGATCAGGTTGAGG	3781
Db	4053	TCCTCAATTTGAAGTCCAAATATTTGCTGTGGGAAGATTTCCAGGATCAGGTTGAGG	4112
Qy	3782	ATTATGCAATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCTTGGACCCATTT	3841
Db	4113	ATTATGCAATTTGAGGAAGAACATATCTGTGCTGAGGTTGAGAAATGCTTGGACCCATTT	4172
Qy	3842	TGGGATATGATACAGACTAATTTTTTTTTTTTTTTT 3875	
Db	4173	TGGGATATGATACAGACTAATTTTTTTTTTTTTTT 4206	

RESULT 12  
AX930571  
LOCUS AX930571 3798 bp DNA linear PAT 22-DEC-2003  
DEFINITION Sequence 43 from Patent WO03087386.  
ACCESSION AX930571  
VERSION AX930571.1 GI:40312319  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Hominidae; Homo.

REFERENCE 1  
AUTHORS Kroeger, B., Zelder, O., Klopffrogge, C., Schroeder, H. and Haefner, S.  
TITLE Methods for producing sulphurous fine chemicals  
JOURNAL Patent: WO 03087386-A 43 23-OCT-2003;  
BASF AKTIENGESSELLSCHAFT (DE)  
FEATURES  
Location/Qualifiers  
1..3798  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1..3798  
/note="unnamed protein product; RHS24705"  
/codon\_start=1  
/protein\_id="CAFO3984.1"  
/db\_xref="GI:40312320"  
translation="MSPALQDLSPQEGKLKTLRDEINAILQKRIWLDGGMGTMLQRE  
KLNHFQGFQEFKDFHRLPKGNNDILSTQPDVYIYIHKYLLAGADIETNTPFSTS  
IAQDYGLEHAYRNMCNSAGVARAAEVLQGTGKRFVAGALGTPVTSVSVE  
RPDYNITFDLVEAYOQAKGLDGGVILLIETIFDTANAKAALFALONFEKYA  
PPPIFGTIVDKSGRTLSGGTGEFVLSVSHGEPLCLGLNCALGAEMRPFIILIGK  
CTTAVLCIPNAGLPNTFGDDETSMAKHLKDFAMDGLVNIVGCCSGCTPDHIREI  
AEAVNCKPRVPPATAFEGHLLSLGLEPFRIGPTNFVNIGERCNVGRKFAKLIMA  
GNYEBALCVAKYQVEMGAQVLDVNDMDGMLDGPSTMRFCNLIASEPDIKVPCLIDS  
SNFVIEAGLKCCKQKCIVNSISLKEGDDFLKXAKIKKYGAANLVAFDEGQATE  
TDTKRVCTRAVHLVVKLGFNPNDIIFDPNLTITGTGMEHNLVAINFIHATKVIKE  
TLPGARISSGLSNLSFSPRGMEAIRAMHGVEFLYHAIKSGMDMGIVNAGNLPYDDIH  
KELLQUCEDLIWNKDPKATEKLLRYAQGTGKVKVITQDEWRNGVEERLEVALVKG  
IEKHIIDTEEARLNQKYPPLNIIEGPLMNGKIVGDLFGAGKMFQVTKSARVM  
KKAVALHLPFEKEREETVNLGTVVEEDPYQGTIVLATVKGVDHIDIGNIVGVLCDS  
NNFRITDLGWTCPDKILKAALDHKADIIGLSGLITPSLDEMI FVAKEMERLAIRPL  
LFGGATSTKWTAVKIAPRYSAPIVHLVDASKSVVVCSSOLDENLKDVEFEEMERVE  
DIRQHYSLKERRIYPLSQARKSGFQMDLSEPHVPKPTFTGTQVFEYDYLQKLVY  
IDWKFFDVQWLRQKYPNRPFKI FNDKTVGGEARKVYDDAHNMLNTLISQKRLRAG  
VFPWPAQSIQDDIHLAYAAVPOAAEPIATFYLQRAEKDASSTPEYICLSDFIAP  
LHSGRDYVLGFAVAFGVEELSKAYEDDDYSIMVKALGDRLEAFAEELHERVR  
RELWAVCGSEOLDVADRLRLRYKIRPAPGYPSPQDPHTKLTWRLADIEQSTGIRLT  
ESLWAPASAVSGLYFSLNLSKYFAVGKISKDQVEDYALRKNI SVAEVEKMLGPLIGY  
DTD"

ORIGIN

Query Match	86.5%;	Score 3390;	DB 6;	Length 3798;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 3790;	Conservative	0;	Mismatches	8;
			Indels	0;
			Gaps	0;
Qy	64	ATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGTCCTGAAGAAAACCCCTGCGGGAT	123	
Db	1	ATGTCAACCCGCGCTCCAGACCTGTGCGAACCCGAGGTCCTGAAGAAAACCCCTGCGGGAT	60	
Qy	124	GAGATCAATGCCATCTTCGAGAGAGGATATGCTGCTGGATGGAGGATGGGACCATG	183	
Db	61	GAGATCAATGCCATCTTCGAGAGAGGATATGCTGCTGGATGGAGGATGGGACCATG	120	
Qy	184	ATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	243	
Db	121	ATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	180	
Qy	244	AGGCCGCTGAAAGGCAACCAATGACATTTTAAGTATAACTCAGCCTGATCTCATTTACCAA	303	
Db	181	AGGCCGCTGAAAGGCAACCAATGACATTTTAAGTATAACTCAGCCTGATCTCATTTACCAA	240	
Qy	304	ATCCATAAGGAATATCTGCTGCTGGGCGAGATATCATTTGAAACAATACTTTTAGCAGC	363	
Db	241	ATCCATAAGGAATATCTGCTGCTGGGCGAGATATCATTTGAAACAATACTTTTAGCAGC	300	
Qy	364	ACTAGTATTTGCCACAGCTGACTATGCTTGAACACTTGCCTTACCGGATGAACATGTGC	423	
Db	301	ACTAGTATTTGCCACAGCTGACTATGCTTGAACACTTGCCTTACCGGATGAACATGTGC	360	
Qy	424	TCTGACGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGAGG	483	
Db	361	TCTGACGAGTGGCCAGAAAAGCTGCCGAGGAGGTAACTCTCCAGACAGGAATTAAGAGG	420	
Qy	484	TTTGTGGCAGGGGCTCTGGGTCGCACTAATAAGACACTCTCTGTGTGCTGCCATCTGTGGAA	543	

Db 421 TTTGTGGAGGGGCTCTGGGTCCGACTTAATAGACACTCTCTGTGTCCCCATCTGTGGAA 480  
Qy 544 AGGCCGGAATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACGAGCC 603  
Db 481 AGGCCGGAATATAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACGAGCC 540  
Qy 604 AAAGGACTTCTGGATGCGGGGTTGATATCTTACTCATTTGAACATAATTTTGTACTGCG 663  
Db 541 AAAGGACTTCTGGATGCGGGGTTGATATCTTACTCATTTGAACATAATTTTGTACTGCG 600  
Qy 664 AATGCCAAGCAGGCTTTGTTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG 723  
Db 601 AATGCCAAGCAGGCTTTGTTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG 660  
Qy 724 CCTATCTTTATTTTCAGGAGCATCGTTGATAAAGTGGCGGACTCTTTTCCGACAGACA 783  
Db 661 CCTATCTTTATTTTCAGGAGCATCGTTGATAAAGTGGCGGACTCTTTTCCGACAGACA 720  
Qy 784 GGAGAGGGATTTGTCATCAGGCTGTCTCATGGAGAACCACTCTGCAATGGATTAATTTGT 843  
Db 721 GGAGAGGGATTTGTCATCAGGCTGTCTCATGGAGAACCACTCTGCAATGGATTAATTTGT 780  
Qy 844 GCTTTGGGTGAGCTGAGATGAGACCTTTTATTCAAATAAATTTGAAATGTTACACAGCC 903  
Db 781 GCTTTGGGTGAGCTGAAATGAGACCTTTTATTTGAAATTAATTTGAAATGTTACACAGCC 840  
Qy 904 TATGTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTTGGTGACTATGATGAACG 963  
Db 841 TATGTCTCTGTTATCCCAATGCAAGGTCTTCCCAACACCTTTTGGTGACTATGATGAACG 900  
Qy 964 CTTTCTATGATGGCCAGCACCTAAAGATTTTGTCTATGATGGCTTGGTCAATATAGTT 1023  
Db 901 CTTTCTATGATGGCCAGCACCTAAAGATTTTGTCTATGATGGCTTGGTCAATATAGTT 960  
Qy 1024 GGAGGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAT 1083  
Db 961 GGAGGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAAT 1020  
Qy 1084 TGTAAAGCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTA 1143  
Db 1021 TGTAAAGCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTA 1080  
Qy 1144 GAGCCCTCAGGATTTGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAAATGTT 1203  
Db 1081 GAGCCCTCAGGATTTGACCGTACACCACTTTGTTTAACTTTGAGAGCGCTGTAAATGTT 1140  
Qy 1204 GCAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACTATGAAGAGCCTTGTGT 1263  
Db 1141 GCAGGATCAAGGAAGTTTGTCTAACTCATATGCGAGGAACTATGAAGAGCCTTGTGT 1200  
Qy 1264 GTTGCCAAAGTGCAGGTGGAAATGGAGCCAGGTGTTGGATGTCAACATGGATGATGGC 1323  
Db 1201 GTTGCCAAAGTGCAGGTGGAAATGGAGCCAGGTGTTGGATGTCAACATGGATGATGGC 1260  
Qy 1324 ATGCTAGATGCTCMAAGTGCAATGACCGAGATTTTGAACCTTAATTTGCTTCGAGCCAGAC 1383  
Db 1261 ATGCTAGATGCTCMAAGTGCAATGACCGAGATTTTGAACCTTAATTTGCTTCGAGCCAGAC 1320  
Qy 1384 ATGCGAAAGGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGTTA 1443  
Db 1321 ATGCGAAAGGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGTTA 1380  
Qy 1444 AAGTCTGCCAAGGGAAGTGCATTTGTCAATAGCAATAGTCTGAAGGAAGGAGAGCGAC 1503  
Db 1381 AAGTCTGCCAAGGGAAGTGCATTTGTCAATAGCAATAGTCTGAAGGAAGGAGAGCGAC 1440  
Qy 1504 TTCTTGGAGAGGCCAGGAAGATTTAAAGATATGAGCTGCTATGCTGGTCAATGCTTTT 1563  
Db 1441 TTCTTGGAGAGGCCAGGAAGATTTAAAGATATGAGCTGCTATGCTGGTCAATGCTTTT 1500  
Qy 1564 GATGAAGAGGACAGGCAACAGAACAGACACAAATTCAGAGTGTGCAACCGCGGCTTAC 1623

Db 1501 GATGAAGAGGACAGGCAACAGAAACAGACACAAATAATCAGAGTGTGACCCGGGCGCTAC 1560  
Qy 1624 CATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTAATTTTTCACCTTAATATC 1683  
Db 1561 CATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTAATTTTTCACCTTAATATC 1620  
Qy 1684 CTAACCAATTTGGGACTGGAATGGAGGAACAACTTGTATGCAATTAATTTTATCCATGCA 1743  
Db 1621 CTAACCAATTTGGGACTGGAATGGAGGAACAACTTGTATGCAATTAATTTTATCCATGCA 1680  
Qy 1744 ACAAAAGTCAATTAAGAAAACTTTACCTGGAGCCAGAAATAGTGAGAGTCTTTTCAACTTG 1803  
Db 1681 ACAAAAGTCAATTAAGAAAACTTTACCTGGAGCCAGAAATAGTGAGAGTCTTTTCAACTTG 1740  
Qy 1804 TCCCTTCTCTTCGAGGAATGGAAGCAATTCGAGAGCAATGCATGGGGTTCCTTTTAC 1863  
Db 1741 TCCCTTCTCTTCGAGGAATGGAAGCAATTCGAGAGCAATGCATGGGGTTCCTTTTAC 1800  
Qy 1864 CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAAACTCCCTGTGTAT 1923  
Db 1801 CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAAACTCCCTGTGTAT 1860  
Qy 1924 GATGATATCCATAAGGAACCTTCTGCAAGTCTGTGAAGATCTCATCTGGAAATAAAGACCT 1983  
Db 1861 GATGATATCCATAAGGAACCTTCTGCAAGTCTGTGAAGATCTCATCTGGAAATAAAGACCT 1920  
Qy 1984 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCCACAGGAGGGAAGAAATC 2043  
Db 1921 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCCACAGGAGGGAAGAAATC 1980  
Qy 2044 ATTCAGACTGATGAGTGGAGAAATGCCCTGTGGAAGAACGCTTCAAGTATGCTTGTG 2103  
Db 1981 ATTCAGACTGATGAGTGGAGAAATGCCCTGTGGAAGAACGCTTCAAGTATGCTTGTG 2040  
Qy 2104 AAGGGCATTTGAAAAACATATTTATGAGGATCTGAGGAAAGCCAGGTAAACCAAAAAA 2163  
Db 2041 AAGGGCATTTGAAAAACATATTTATGAGGATCTGAGGAAAGCCAGGTAAACCAAAAAA 2100  
Qy 2164 TATCCCGGACCTCTCAATATAATTTGAAGAACCTCTGATGAATGGATGAATTTGTTGT 2223  
Db 2101 TATCCCGGACCTCTCAATATAATTTGAAGAACCTCTGATGAATGGATGAATTTGTTGT 2160  
Qy 2224 GATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAGCCGGGTATG 2283  
Db 2161 GATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAGCCGGGTATG 2220  
Qy 2284 AAGAAGCTGTGCGCACTTATCCCTTTTCATGGAATAAAGAGAGAAACACAGATG 2343  
Db 2221 AAGAAGCTGTGCGCACTTATCCCTTTTCATGGAATAAAGAGAGAAACACAGATG 2280  
Qy 2344 CTTAAACGGCAGTAGAAGAGGAGGACCTTTACAGGGCACCATCTGCTGGCCACTGTT 2403  
Db 2281 CTTAAACGGCAGTAGAAGAGGAGGACCTTTACAGGGCACCATCTGCTGGCCACTGTT 2340  
Qy 2404 AAAGGGACCTGACGACATAGCAGAAACATAGTTGGAGTAGTCTTGGCTGCAATAAT 2463  
Db 2341 AAAGGGACCTGACGACATAGCAGAAACATAGTTGGAGTAGTCTTGGCTGCAATAAT 2400  
Qy 2464 TTCCGAGTTATGATTTAGGAGTCAAGCTCCATGTTGATTAAGATACGTAAAGCTGCTCT 2523  
Db 2401 TTCCGAGTTATGATTTAGGAGTCAAGCTCCATGTTGATTAAGATACGTAAAGCTGCTCT 2460  
Qy 2524 GACCAAAAGCAGATATAATTTGGCTGTCAAGACTCATCTCTTCCCTGGATGAATG 2583  
Db 2461 GACCACAAAGCAGATATAATTTGGCTGTCAAGACTCATCTCTTCCCTGGATGAATG 2520  
Qy 2584 ATTTTGTGTCGAAGGAATGGAGAGATTAGCTATAAGGATTCCATTTGTTGATTTGGAGGA 2643  
Db 2521 ATTTTGTGTCGAAGGAATGGAGAGATTAGCTATAAGGATTCCATTTGTTGATTTGGAGGA 2580  
Qy 2644 GCAACCACTTCAAAAAACCCACACAGCAGGATTTAAATAGCTCCGAGATACAGTGCACTGTA 2703  
Db 2581 GCAACCACTTCAAAAAACCCACACAGCAGGATTTAAATAGCTCCGAGATACAGTGCACTGTA 2640





Db 873 AAAGCCGGATTATAGGAACATCACATTTGATGAGCTTGTGTAAGCATACCAAGAGCAGG 932  
Qy 602 CCAAAGGACTTCTGGATGGGAGGTTGATATCTTATCTCATTTGAAACCTATTTTGTGATCTG 661  
Db 933 CCAAAGGACTTCTGGATGGGAGGTTGATATCTTACTCATTTGAAACCTATTTTGTGATCTG 992  
Qy 662 CCAATGCCAAGGAGCCTTGTGTCATCCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
Db 993 CCAATGCCAAGGAGCCTTGTGTCATCCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTTTCAAGGACGATCGTTGATATAAAGTGGGCGGACTCTTTTCCGGACAGA 781  
Db 1053 GGCCTATCTTTATTTCAAGGAGGATCGTTGATATAAAGTGGGCGGACTCTTTTCCGGACAGA 1112  
Qy 782 CAGGAGAGGATTTGTCTCATCAGCGTGTCTCATGAGGAACCACTCTCTGCAATTTGGAATTTAAAT 841  
Db 1113 CAGGAGAGGATTTGTCTCATCAGCCTGTCTCATGAGGAACCACTCTCTGCAATTTGGAATTTAAAT 1172  
Qy 842 GTGCTTTGGGTGGCAGCTGAGATGAGACCTTTTATTGAAATAAATTTGGAATAATGTACAACAG 901  
Db 1173 GTGCTTTGGGTGGCAGCTGAAATGAGACCTTTTATTGAAATAAATTTGGAATAATGTACAACAG 1232  
Qy 902 CCTATGTCCTCTGTTATCCCAATGCAGGTCTTCCCAACACCTTTTGGTGACTATGATGAAA 961  
Db 1233 CCTATGTCCTCTGTTATCCCAATGCAGGTCTTCCCAACACCTTTTGGTGACTATGATGAAA 1292  
Qy 962 GGCCTTCTATGATGGCCAAAGCACTTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
Db 1293 GGCCTTCTATGATGGCCAAAGCACCCTTAAAGGATTTTGTCTATGATGGCTTGGTCAATATAG 1352  
Qy 1022 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTTGAGCTPAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGCTCGTC 1141  
Db 1413 ATTTGAGCTPAGAGTTCCACTGCCACTGCTTTTGAAGGACATATGTTACTGCTCGTC 1472  
Qy 1142 TAGAGCCTTCAGGATTTGGACCGTACACAACTTTGTTTAACTTTGGAGAGCGCTGTAATG 1201  
Db 1473 TAGAGCCTTCAGGATTTGGACCGTACACAACTTTGTTTAACTTTGGAGAGCGCTGTAATG 1532  
Qy 1202 TTGCAGGATCAAGGAGTTTGTCTAACTCATCATGCGAGGAAATCATGAAAGCCTTGT 1261  
Db 1533 TTGCAGGATCAAGGAGTTTGTCTAACTCATCATGCGAGGAAATCATGAAAGCCTTGT 1592  
Qy 1262 GTGTTGCCAAAGTGCAGGTGAAATGGAGGCCAGGTGTTGGATGTCAAATGATGATG 1321  
Db 1593 GTGTTGCCAAAGTGCAGGTGAAATGGAGGCCAGGTGTTGGATGTCAAATGATGATG 1652  
Qy 1322 GCATGCTAGATGGTCCAAAGTGCATGACCAAGATTTTGCAACTTTAATTTGCTTCCGAGCAG 1381  
Db 1653 GCATGCTAGATGGTCCAAAGTGCATGACCAAGATTTTGCAACTTTAATTTGCTTCCGAGCAG 1712  
Qy 1382 ACATGCGAAAGGTACTTTGTGCTGCTGCTCTCCAAATTTTGTGATGTAAGCTGGGT 1441  
Db 1713 ACATGCGAAAGGTACTTTGTGCTGCTGCTCTCCAAATTTTGTGATGTAAGCTGGGT 1772  
Qy 1442 TAAAGTGTGCGCAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCAG 1501  
Db 1773 TAAAGTGTGCGCAAGGGAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAAGGAGGAGCAG 1832  
Qy 1502 ACTTCTTGGAGAGGCCAGGGAAGTATAAAGTATGGAGCTGCTATGTTGTTGATGCTGCTT 1561  
Db 1833 ACTTCTTGGAGAGGCCAGGGAAGTATAAAGTATGGAGCTGCTATGTTGTTGATGCTGCTT 1892  
Qy 1562 TTGATGAGAGGACAGGCAACAGAAACAGACAAATCAGAGTGTGCAACCGGCT 1621  
Db 1893 TTGATGAGAGGACAGGCAACAGAAACAGACAAATCAGAGTGTGCAACCGGCT 1952  
Qy 1622 ACCATCTGCTGTGAAAAAATGGGCTTTAATCCAAATGACATTTATTTTTCACCTTAATA 1681  
Db 1953 ACCATCTGCTGTGAAAAAATGGGCTTTAATCCAAATGACATTTATTTTTCACCTTAATA 2012

Qy 1682 TCCTAACCAATTGGGACTGGAATGGAGGAACAACTTGTGATGCCATTAATTTTATCCATG 1741  
Db 2013 TCCTAACCAATTGGGACTGGAATGGAGGAACAACTTGTGATGCCATTAATTTTATCCATG 2072  
Qy 1742 CAAACAAAGTCAATTAAGAAAACTTTACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 1801  
Db 2073 CAAACAAAGTCAATTAAGAAAACTTTACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 2132  
Qy 1802 TGTCTTCTTCTTCCGAGGAATGGAAAGCCATTTGCGAAGCAATGCAATGGGTTTCTCTTT 1861  
Db 2133 TGTCTTCTTCTTCCGAGGAATGGAAAGCCATTTGCGAAGCAATGCAATGGGTTTCTCTTT 2192  
Qy 1862 ACCATGCAATCAAGTCTGGCATGGAATAGTGAATGCTGGAACCTCTCCCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGGCATGGAATAGTGAATGCTGGAACCTCTCCCTGTGT 2252  
Qy 1922 ATGATGATATCCATAAGGAACCTTTGCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db 2253 ATGATGATATCCATAAGGAACCTTTGCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
Qy 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCACAGGAGGGAAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCACAGGAGGGAAGAAAG 2372  
Qy 2042 TCATTCAGACTGATGATGAGGAAATGGGCCCTGTGCGAAGAACGCTTGGAGTATGCCCTTG 2101  
Db 2373 TCATTCAGACTGATGATGAGGAAATGGGCCCTGTGCGAAGAACGCTTGGAGTATGCCCTTG 2432  
Qy 2102 TGAAGGGCAATGAAAAACATATTTATTGAGGATCTGAGGAAGCCAGGTTAAACCAAAAAA 2161  
Db 2433 TGAAGGGCAATGAAAAACATATTTATTGAGGATCTGAGGAGCCAGGTTAAACCAAAAAA 2492  
Qy 2162 AATATCCCGACCTCTCAATATATAATTGAAGGACCCCTGATGAAATGGAATTTGTTG 2221  
Db 2493 AATATCCCGACCTCTCAATATATAATTGAAGGACCCCTGATGAAATGGAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAATAAATCTTCTACTCAGGTTATAAGTTCAGCCCGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAATAAATCTTCTACTCAGGTTATAAGTTCAGCCCGGTTA 2612  
Qy 2282 TGAAGAGGCTGTGTGGCCACTTATCCCTTTTCATGGAAGAAAGAGAGAAACACAGAG 2341  
Db 2613 TGAAGAGGCTGTGTGGCCACTTATCCCTTTTCATGGAAGAAAGAGAGAAACACAGAG 2672  
Qy 2342 TGCTTAAACGCGCACAGTAGAAGAGAGGACCCCTTACAGGGCACCATGCTGCGCCACTG 2401  
Db 2673 TGCTTAAACGCGCACAGTAGAAGAGAGGACCCCTTACAGGGCACCATGCTGCGCCACTG 2732  
Qy 2402 TTAAGGCGAGCGTGCACGATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2461  
Db 2733 TTAAGGCGAGCGTGCACGATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2792  
Qy 2462 ATTTCCGAGTTATTGATTTAGGAGTCAATGACTCCATGTAAGATACTGAAAGCTGCTC 2521  
Db 2793 ATTTCCGAGTTATTGATTTAGGAGTCAATGACTCCATGTAAGATACTGAAAGCTGCTC 2852  
Qy 2522 TTGAACCAAAAGCAGATATAATTTGGCTGTCAAGACTCATCACTCTCTTCCCTGGATGAAA 2581  
Db 2853 TTGAACCAAAAGCAGATATAATTTGGCTGTCAAGACTCATCACTCTCTTCCCTGGATGAAA 2912  
Qy 2582 TGATTTTGTGTCGAAGGAATGGAGAGATTAGCTATAAGGATTCATTTGATTCGAG 2641  
Db 2913 TGATTTTGTGTCGAAGGAATGGAGAGATTAGCTATAAGGATTCATTTGATTCGAG 2972  
Qy 2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACCTG 2701  
Db 2973 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACCTG 3032  
Qy 2702 TAATPCTATGCTCTGGAGCGGCTTCAAGAGTGTGTTGTTTCCAGCTGTTAGATGAAA 2761  
Db 3033 TAATPCTATGCTCTGGAGCGGCTTCAAGAGTGTGTTGTTTCCAGCTGTTAGATGAAA 3092



QY 2762 ATCTAAGGATGATATCTTTGAGGAATCATGGAAGATATGAAGATATTTAGACAGACC 2821  
Db 3093 ATCTAAGGATGATATCTTTGAGGAATCATGGAAGATATTTAGACAGACC 3152  
QY 2822 ATTATGAGTCTCTCAAGGAGAGAGATCTTACCCCTTAAGTCAAGCCAGAAAAGTGGTT 2881  
Db 3153 ATTATGAGTCTCTCAAGGAGAGAGATCTTACCCCTTAAGTCAAGCCAGAAAAGTGGTT 3212  
QY 2882 TCCAAATGGATGGCTGTCTGAACCTCACCCAGTGAAGCCACCGTTTATTGGACCCAGG 2941  
Db 3213 TCCAAATGGATGGCTGTCTGAACCTCACCCAGTGAAGCCACCGTTTATTGGACCCAGG 3272  
QY 2942 TCTTTGAAGACTATGACCTGCAGAACTGCTGGAAGTCTGACTACATTTGAGAGCCCTTCTTTG 3001  
Db 3273 TCTTTGAAGACTATGACCTGCAGAACTGCTGGAAGTCTGACTACATTTGAGAGCCCTTCTTTG 3332  
QY 3002 ATGCTCGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTCCCAAGATATTTAAGACCA 3061  
Db 3333 ATGCTCGGAGCTCCGGGGCAAGTACCCGAATCAGAGCTTCCCAAGATATTTAAGACCA 3392  
QY 3062 AAACAGTAGGTGAGAGGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3121  
Db 3393 AAACAGTAGGTGAGAGGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3452  
QY 3122 TGATTAGTCAAAAGAACTCCGGGGCCGGGTGTGGTTCTGGCCAGACAGAGTA 3181  
Db 3453 TGATTAGTCAAAAGAACTCCGGGGCCGGGTGTGGTTCTGGCCAGACAGAGTA 3512  
QY 3182 TCCAAAGACGATTTCACTCTGACGAGAGGCTGCTGTGCCCGCCAGGCTTGACAGCCCATAG 3241  
Db 3513 TCCAAAGACGATTTCACTCTGACGAGAGGCTGCTGTGCCCGCCAGGCTTGACAGCCCATAG 3572  
QY 3242 CCACTTTCTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3301  
Db 3573 CCACTTTCTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3632  
QY 3302 ACTGCCCTCTCAGACTTCATCGCTCCCTTGCAATCTGCGATCCCTGACTACCTGGGCTGT 3361  
Db 3633 ACTGCCCTCTCAGACTTCATCGCTCCCTTGCAATCTGCGATCCCTGACTACCTGGGCTGT 3692  
QY 3362 TTCCCGTTGCTCTCTTTGGGGTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTAGC 3421  
Db 3693 TTCCCGTTGCTCTCTTTGGGGTAGAGAGCTGAGCAAGGCTATGAGGATGATGGTAGC 3752  
QY 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAGAG 3481  
Db 3753 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCTTTGCAAGAG 3812  
QY 3482 AGCTCCATGAAGAGTTCCGCGGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGAG 3541  
Db 3813 AGCTCCATGAAGAGTTCCGCGGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGAG 3872  
QY 3542 TCCAGACCTGCGAAGTTTCCGCTACAGAGGCTCCGCGGCTCCCTGGCTACCCAGCC 3601  
Db 3873 TCCAGACCTGCGAAGTTTCCGCTACAGAGGCTCCGCGGCTCCCTGGCTACCCAGCC 3932  
QY 3602 AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
Db 3933 AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3992  
QY 3662 GCATTAGGTTAAAGAAATATTAGCAATGCGACTGCTGCGAGTCTCAGGCTCTACT 3721  
Db 3993 GCATTAGGTTAAAGAAATATTAGCAATGCGACTGCTGCGAGTCTCAGGCTCTACT 4052  
QY 3722 TCTCCATTTGAGTCCAAATATTGCTGCTGGGAGAGATTTCCAGGATCAGGTTGAG 3781  
Db 4053 TCTCCATTTGAGTCCAAATATTGCTGCTGGGAGAGATTTCCAGGATCAGGTTGAG 4112  
QY 3782 ATTATGATTGAGGAAGCAATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
Db 4113 ATTATGATTGAGGAAGCAATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4172  
QY 3842 TGGGATATGATACAGACTAATTTTTTTTTTTTTTTT 3875

Db 4173 TGGGATATGATACAGACTAATTTTTTTTTTTTTTTT 4206  
RESULT 14  
LOCUS AR438495 7224 bp DNA linear PAT 20-FEB-2004  
DEFINITION Sequence 1 from patent US 6664062.  
ACCESSION AR438495  
VERSION AR438495.1 GI:42663366  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS Stanton,V.P. Jr.  
TITLE Thymidylate synthase gene sequence variances having utility in  
determining the treatment of disease  
JOURNAL Patent: US 6664062-A 1 16-DEC-2003;  
Nuvelo, Inc.; Sunnyvale, CA  
FEATURES  
Location/Qualifiers  
1..7224  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 79.3%; Score 3109; DB 6; Length 7224;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3859; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 2 GTCACTGTGAGAGACGCTCTTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 61  
Db 333 GTCACTGTGAGAGACGCTCTTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGACA 392  
QY 62 ACATGTCACTCCCGCTCCAAAGACCTGTCGAACCCGAGGCTCTGAAGAAAACCTCGCGG 121  
Db 393 ACATGTCACTCCCGCTCCAAAGACCTGTCGAACCCGAGGCTCTGAAGAAAACCTCGCGG 452  
QY 122 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTATGCTGCTGGATGGAGGATGGGACCA 181  
Db 453 ATGAGATCAATGCCATTTCTGCAGAAAGAGATTATGCTGCTGGATGGAGGATGGGACCA 512  
QY 182 TGATTCAGCGGAGAGAGCTTAACGAGAAACACTTTCGAGGTCAGGAATTTAAGATCATG 241  
Db 513 TGATTCAGCGGAGAGAGCTTAACGAGAAACACTTTCGAGGTCAGGAATTTAAGATCATG 572  
QY 242 CCAGCGCTGTAAGAGCAACATGACATTTTAAGTATTAATCACTCAGCCTGATGTCATTTACC 301  
Db 573 CCAGCGCTGTAAGAGCAACATGACATTTTAAGTATTAATCACTCAGCCTGATGTCATTTACC 632  
QY 302 AAATCCATAAGGAATACTTTGCTGGCTGGGCGCAGATATCAATTGAAAACAAATACCTTTAGCA 361  
Db 633 AAATCCATAAGGAATACTTTGCTGGCTGGGCGCAGATATCAATTGAAAACAAATACCTTTAGCA 692  
QY 362 GCATAGTATTGCCAAGCTGACTATGCGCTTGAACACTTGGCCTTACCGGATGAACATGT 421  
Db 693 GCATAGTATTGCCAAGCTGACTATGCGCTTGAACACTTGGCCTTACCGGATGAACATGT 752  
QY 422 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTTAAGA 481  
Db 753 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTTAAGA 812  
QY 482 GGTGTTGTGCAGGGGCTCTGGGCTCCGACTTAATAAGACACTCTCTGTGTCCTCATCTGTGG 541  
Db 813 GGTGTTGTGCAGGGGCTCTGGGCTCCGACTTAATAAGACACTCTCTGTGTCCTCATCTGTGG 872  
QY 542 AAAGCCCGGATTTATAGGAACATCACTTGTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
Db 873 AAAGCCCGGATTTATAGGAACATCACTTGTGATGAGCTTGTGAAGCATACCAAGAGCAGG 932  
QY 602 CCAAGGAGCTTCTGGATGGCGGGGTTGATCTTACTCATTTGAAGACTATTTTGTACTG 661  
Db 933 CCAAGGAGCTTCTGGATGGCGGGGTTGATCTTACTCATTTGAAGACTATTTTGTACTG 992

QY 662 CCAATGCCAAGGAGCCTTCTTGGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
DB 993 CCAATGCCAAGGAGCCTTCTTGGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052  
QY 722 GGCTATCTTTATTTCAAGGACGATCGTTGATATAAAGTGGCGGAGCTCTTTCCGGACAGA 781  
DB 1053 GGCTATCTTTATTTCAAGGACGATCGTTGATATAAAGTGGCGGAGCTCTTTCCGGACAGA 1112  
QY 782 CAGGAGAGGATTTGTTCATCAGCGGTCTCATGGAGAACCACTCTCGATTTGGATTAAAT 841  
DB 1113 CAGGAGAGGATTTGTTCATCAGCGGTCTCATGGAGAACCACTCTCGATTTGGATTAAAT 1172  
QY 842 GTGCTTTGGGTGACGTGACATGAGACCTTTTATTTGAATAAATTTGGAAATGTACACAG 901  
DB 1173 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTTGAATAAATTTGGAAATGTACACAG 1232  
QY 902 CCTATGCTCTGTATATCCCAATGACGCTCTCCCAACACCTTTGGTGAATATGATGAAA 961  
DB 1233 CCTATGCTCTGTATATCCCAATGACGCTCTCCCAACACCTTTGGTGAATATGATGAAA 1292  
QY 962 CGCCTTCTATGATGGCCAGCACTAAAGATTTTGTATGGATGGCTTGGTCAATATAG 1021  
DB 1293 CGCCTTCTATGATGGCCAGCACTAAAGATTTTGTATGGATGGCTTGGTCAATATAG 1352  
QY 1022 TTGGAGGATGCTTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
DB 1353 TTGGAGGATGCTTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
QY 1082 ATTGTAAGCTTAGAGTTCCACCTGCCACTGCTTTGAAGACACATATGTAATGCTGTGTC 1441  
DB 1413 ATTGTAAGCTTAGAGTTCCACCTGCCACTGCTTTGAAGACACATATGTAATGCTGTGTC 1472  
QY 1442 TAGAGCCTTCAGATTTGAGACCTGACCAACTTTTGTATTAACATTTGGAGAGCGCTGTAATG 1201  
DB 1473 TAGAGCCTTCAGATTTGAGACCTGACCAACTTTTGTATTAACATTTGGAGAGCGCTGTAATG 1532  
QY 1201 TTGCAGGATCAAGAACTTTGCTTAACTCATATGCGAGGAACTATGAAGAGCCTTGT 1261  
DB 1533 TTGCAGGATCAAGAACTTTGCTTAACTCATATGCGAGGAACTATGAAGAGCCTTGT 1592  
QY 1262 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCAGCTGTTGGATGTCAACATGGATGATG 1321  
DB 1593 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCAGCTGTTGGATGTCAACATGGATGATG 1652  
QY 1322 GCATGCTAGATGGTCCAAAGTGCATATGACCAAGATTTTGCATTTAATTTCCAGCCAG 1381  
DB 1653 GCATGCTAGATGGTCCAAAGTGCATATGACCAAGATTTTGCATTTAATTTCCAGCCAG 1712  
QY 1382 ACATGCGCAAGGTACCTTTGTGCATCGACTCTCCATTTTGTGCTGATTTGAAGCTGGT 1441  
DB 1713 ACATGCGCAAGGTACCTTTGTGCATCGACTCTCCATTTTGTGCTGATTTGAAGCTGGT 1772  
QY 1442 TAAAGTCTGCCAAGGAGTGCAATGTCAATAGCATTTAGTCTCAAGGAAGAGAGACG 1501  
DB 1773 TAAAGTCTGCCAAGGAGTGCAATGTCAATAGCATTTAGTCTCAAGGAAGAGAGACG 1832  
QY 1502 ACTTCTTTGGAGAGCCAGGAGATTTAAAGTATGGAGCTGCTATGGTGTCAATGCTT 1561  
DB 1833 ACTTCTTTGGAGAGCCAGGAGATTTAAAGTATGGAGCTGCTATGGTGTCAATGCTT 1892  
QY 1562 TTGATGAAGAGGACAGGCAACAGAACAGACACAAAATCAGAGTGTGCAACCGGCGCT 1621  
DB 1893 TTGATGAAGAGGACAGGCAACAGAACAGACACAAAATCAGAGTGTGCAACCGGCGCT 1952  
QY 1622 ACCATCTGCTTGTGAAAAAATGCGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1681  
DB 1953 ACCATCTGCTTGTGAAAAAATGCGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 2012  
QY 1682 TCCTAACCAATTTGGGACTGGAATGGAGGAAACAACTTGTATGCGCAATTAATTTTATCCATG 1741  
DB 2013 TCCTAACCAATTTGGGACTGGAATGGAGGAAACAACTTGTATGCGCAATTAATTTTATCCATG 2072

QY 1742 CAACAAAAGTCATTAAAGAAAACATTAACCTGGAGCCAGATAAAGTGAGGAGTCTTTTCCAACT 1801  
DB 2073 CAACAAAAGTCATTAAAGAAAACATTAACCTGGAGCCAGATAAAGTGAGGAGTCTTTTCCAACT 2132  
QY 1802 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTGGAAGCAATGCAATGGGGTTTTCCTTT 1861  
DB 2133 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTGGAAGCAATGCAATGGGGTTTTCCTTT 2192  
QY 1862 ACCATGCAATCAAGTCTGGCATGGACATGAGATAGTGAATCTGGAACCTTCCCTGTGT 1921  
DB 2193 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATCTGGAACCTTCCCTGTGT 2252  
QY 1922 ATGATGATATCATTAAGGAACCTTTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
DB 2253 ATGATGATATCATTAAGGAACCTTTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
QY 1982 CTGAGGCCACTGAGAAAGCTTTAOGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2041  
DB 2313 CTGAGGCCACTGAGAAAGCTTTAOGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2372  
QY 2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAGCCCTTGAGTATGCCCTTG 2101  
DB 2373 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTCGAAGAGCCCTTGAGTATGCCCTTG 2432  
QY 2102 TGAAGGGCAATGAAAAACATATTTATTTGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA 2161  
DB 2433 TGAAGGGCAATGAAAAACATATTTATTTGAGGATCTGAGGAAGCCAGGTTAAACCAAAAA 2492  
QY 2162 AATATCCCAGACCTCTCAATATATAATTTGAAGGACCCCTGATGAAATGGAATGAAATTTGTTG 2221  
DB 2493 AATATCCCAGACCTCTCAATATATAATTTGAAGGACCCCTGATGAAATGGAATGAAATTTGTTG 2552  
QY 2222 GTGATCTTTTGGAGCTGGAAATAATGTTTCTACCTCAGGTTATAAGTTCAGCCCGGTTA 2281  
DB 2553 GTGATCTTTTGGAGCTGGAAATAATGTTTCTACCTCAGGTTATAAGTTCAGCCCGGTTA 2612  
QY 2282 TGAAGGAAGGCTGTGTGCCACCTTTATCTCTTTCATGAAAAAGAAAGAGAAACACAGAG 2341  
DB 2613 TGAAGGAAGGCTGTGTGCCACCTTTATCTCTTTCATGAAAAAGAAAGAGAAACACAGAG 2672  
QY 2342 TGCTTAAACGGCACAGTAGAAGAGGAGCCCTTTACAGGGGACCATCTGCTGGCCACTG 2401  
DB 2673 TGAAGGGCAGCTGTGCACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2461  
QY 2402 TGAAGGGCAGCTGTGCACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2492  
DB 2733 TGAAGGGCAGCTGTGCACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATA 2792  
QY 2462 ATTTCCGAGTTATTTGATTTAGGAGTCACTGACTCCATGTGATAAGATACTGAAAAGCTGCTC 2521  
DB 2793 ATTTCCGAGTTATTTGATTTAGGAGTCACTGACTCCATGTGATAAGATACTGAAAAGCTGCTC 2852  
QY 2522 TTGAACCAAAAGCAGATATAAATTTGGCTGTGAGGACTCATCTCTTCCCTCGATGAAA 2581  
DB 2853 TTGAACCAAAAGCAGATATAAATTTGGCTGTGAGGACTCATCTCTTCCCTCGATGAAA 2912  
QY 2582 TGATTTTGTGTCAGAGAAATGGAAGATTTAGCTATAGGATTCATTTGTTGATTTGGAG 2641  
DB 2913 TGATTTTGTGTCAGAGAAATGGAAGATTTAGCTATAGGATTTCAATTTGTTGATTTGGAG 2972  
QY 2642 GAGCAACCACTTCAAAAAACCAACACAGCAGTAAAAATAGCTCCGAGATACAGTGCACTG 2701  
DB 2973 GAGCAACCACTTCAAAAAACCAACACAGCAGTAAAAATAGCTCCGAGATACAGTGCACTG 3032  
QY 2702 TAATCCATGTCTTGGAGCGCTCCAAGAGTGTGTGTGTTCCTCAGCTGTTAGATGAAA 2761  
DB 3033 TAATCCATGTCTTGGAGCGCTCCAAGAGTGTGTGTGTTCCTCAGCTGTTAGATGAAA 3092  
QY 2762 ATCTAAAGGATGATATCTTTGAGGAATCATGGAAGATATGAGATATTAGACAGGACC 2821  
DB 3093 ATCTAAAGGATGATATCTTTGAGGAATCATGGAAGATATGAGATATTAGACAGGACC 3152  
QY 2822 ATTATGAGTCTCTCAAGGAGGAGAGATACCTTACCTTTAAGTCAAGCCAGAAAAAGTGTGTT 2881

```
Db 3153 ATTATGAGTCTCTCAAGAGGAGGATCTTACCTTAACTCAAGCCAGGAAAAANTNGTT 3212
Qy 2882 TCCAAATGGAATGGCTGTCTGAACCTCACCAAGTGAAGCCACGTTTATTGGACCCAGG 2941
Db 3213 TCCAAATGGAATGGCTGTCTGAACCTCACCAAGTGAAGCCACGTTTATTGGACCCAGG 3272
Qy 2942 TCTTTGAAGACTATGACTCGCAGAGCTGGTGGACTACATTTGACTGGAAGCCTTCTTTG 3001
Db 3273 TCTTTGAAGACTATGACTCGCAGAGCTGGTGGACTACATTTGACTGGAAGCCTTCTTTG 3332
Qy 3002 ATGTCTGGCAGCTCCGGGGCAAGTACCAGGCTTCCAGGCTTCCCAAGATATTTAAACACA 3061
Db 3333 ATGTCTGGCAGCTCCGGGGCAAGTACCAGGCTTCCAGGCTTCCCAAGATATTTAAACACA 3392
Qy 3062 AAACAGTAGTGGAGAGGCGAGGAAGTCTACGATGATGCCCAAAATATGCTGAACACAC 3121
Db 3393 AAACAGTAGTGGAGAGGCGAGGAAGTCTACGATGATGCCCAAAATATGCTGAACACAC 3452
Qy 3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTGTGGTCTGGCCAGCAGAGTA 3181
Db 3453 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTGTGGTCTGGCCAGCAGAGTA 3512
Qy 3182 TCCAAAGACGACATTCACCTGTACGACAGAGCTGTGTGCCCGCCAGGCTGCAGAGCCATAG 3241
Db 3513 TCCAAAGACGACATTCACCTGTACGCGGAGGCTGTGTGCCCGCCAGGCTGCAGAGCCATAG 3572
Qy 3242 CCACTTTCTATGGGTTAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGCCATCT 3301
Db 3573 CCACCTTCTATGGTTAAGCAACAGGCTGAGAAGGACTCTGCCAGCAGCGAGCCATCT 3632
Qy 3302 ACTGCCCTCAGACTTCATCGCTCCCTTGTGATTTCTGGCATTCGCTGACTACCTGGGCTGT 3361
Db 3633 ACTGCCCTCAGACTTCATCGCTCCCTTGTGATTTCTGGCATTCGCTGACTACCTGGGCTGT 3692
Qy 3362 TTCCCGTTGCTGCTTTGGGGTAGAGAGCTGAGCAGAGGCTATGAGGATGATGGTGACG 3421
Db 3693 TTCCCGTTGCTGCTTTGGGGTAGAGAGCTGAGCAGAGGCTATGAGGATGATGGTGACG 3752
Qy 3422 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCAGAG 3481
Db 3753 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCAGAG 3812
Qy 3482 AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGAAG 3541
Db 3813 AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGAAG 3872
Qy 3542 TCGCAGACCTCGGAGGTTCCGGTACAGGGGCTCCGCGGCTCCCTGGCTACCCAGCC 3601
Db 3873 TCGCAGACCTCGGAGGCTCGGTTACAGGGGCTCCGCGGCTCCCTGGCTACCCAGCC 3932
Qy 3602 AGCCCGACACACCGAGAACTCACCATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3661
Db 3933 AGCCCGACACACCGAGAACTCACCATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3992
Qy 3662 GCATTAGTTAAAGAAATCATTAGCAATGCGACTGTGCTCAGCAGTCTCAGGCTCTACT 3721
Db 3993 GCATTAGTTAAAGAAATCATTAGCAATGCGACTGTGCTCAGCAGTCTCAGGCTCTACT 4052
Qy 3722 TCTCCAAATTTGAATCCAAATATTTTGTGTGGGGAGATTTCCAGGATCAGGTTGAG 3781
Db 4053 TCTCCAAATTTGAATCCAAATATTTTGTGTGGGGAGATTTCCAGGATCAGGTTGAG 4112
Qy 3782 ATTATGATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGAAGCCATTT 3841
Db 4113 ATTATGATTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGAAGCCATTT 4172
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 3875
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 4206
```

RESULT 15

```
AR562003 LOCUS AR562003 7224 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759200.
ACCESSION AR562003
VERSION AR562003.1 GI:53975655
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7224)
AUTHORS Stanton,V.P., Jr.
TITLE Thymidine phosphorylase gene sequence variances having utility in
determining the treatment of disease
JOURNAL Patent: US 6759200-A 1 06-JUL-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source Location/Qualifiers
1..7224
/mol_type="genomic DNA"
```

Query Match 79.3%; Score 3109; DB 6; Length 7224;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3859; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy 2 GTACCTGTGGAGAGACGCTCTTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 61
Db 333 GTACCTGTGGAGAGACGCTCTTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 392
Qy 62 ACATGTCAACCCCGCTCCAAAGACCTGTGCGAACCCGAGGCTCTGAGAAACCCCTCGGG 121
Db 393 ACATGTCAACCCCGCTCCAAAGACCTGTGCGAACCCGAGGCTCTGAGAAACCCCTCGGG 452
Qy 122 ATGAGATCAATGCATTTCTGCGAAGAGAGATTTATGCTGTGATGAGGAGATGGGACCA 181
Db 453 ATGAGATCAATGCCATTTCTGCGAAGAGAGATTTATGCTGTGATGAGGAGATGGGACCA 512
Qy 182 TGATTCAGCGGAGAGAGCTTAAACGAAAGACACTTCCGAGGTCAGGAATTTAAAGATCATG 241
Db 513 TGATTCAGCGGAGAGAGCTTAAACGAAAGACACTTCCGAGGTCAGGAATTTAAAGATCATG 572
Qy 242 CCAGGCGCTGAAAGGCAACATGATCTTTAAGTATTAATCACTCAGCTGTGATTTACC 301
Db 573 CCAGGCGCTGAAAGGCAACATGATCTTTAAGTATTAATCACTCAGCTGTGATTTACC 632
Qy 302 AAATCCATAAGCAATACTTGTGCTGGGCGCAGATATCATTTGAAACAAATACTTTTAGCA 361
Db 633 AAATCCATAAGCAATACTTGTGCTGGGCGCAGATATCATTTGAAACAAATACTTTTAGCA 692
Qy 362 GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAAACACTTGGCCTACCGGATGAACATGT 421
Db 693 GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAAACACTTGGCCTACCGGATGAACATGT 752
Qy 422 GCTCTGAGAGTGGCGAGAAAAGCTGCGAGAGGTAATCTCTCCAGACAGGAATTAAGA 481
Db 753 GCTCTGAGAGTGGCGAGAAAAGCTGCGAGAGGTAATCTCTCCAGACAGGAATTAAGA 812
Qy 482 GGTTTGTGCGAGGGCTGTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCATCTGTGG 541
Db 813 GGTTTGTGCGAGGGCTGTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCATCTGTGG 872
Qy 542 AAAGGCGGATATAGGAACATCACATTTTGTGAGCTTGTGGAAGCATACCAAGAGCAGG 601
Db 873 AAAGGCGGATATAGGAACATCACATTTTGTGAGCTTGTGGAAGCATACCAAGAGCAGG 932
Qy 602 CCAAGGACTTCTGATGCGGGGTTGATCTTACTCATTTGAATACTATTTTGTACTG 661
Db 933 CCAAGGACTTCTGATGCGGGGTTGATCTTACTCATTTGAATACTATTTTGTACTG 992
Qy 662 CCAATGCCAAGGAGCCTTGTGTGACCTCCAAAATCTTTTTCGAGGAGAAATATGCTCCC 721
Db 993 CCAATGCCAAGGAGCCTTGTGTGACCTCCAAAATCTTTTTCGAGGAGAAATATGCTCCC 1052
```

QY 722 GGCTATCTTTATTTTTCAGGACGATCGTTGATATAAGTGGCGGACCTCTTTCGGACAGA 781  
DB 1053 GGCTATCTTTATTTTTCAGGACGATCGTTGATATAAGTGGCGGACCTCTTTCGGACAGA 1112  
QY 782 CAGGAGAGGATTTGTTCATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGATTTAAAT 841  
DB 1113 CAGGAGAGGATTTGTTCATCAGCWTGTCTCATGAGAACCACTCTGCAATTTGATTTAAAT 1172  
QY 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATAAATTTGAAATAATGTAACAAG 901  
DB 1173 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTTGAAATAAATTTGAAATAATGTAACAAG 1232  
QY 902 CCTATGCTCTGTATTTCCCAATGACAGGTCTTCCCAACACTTTGGTGAATGATGAA 961  
DB 1233 CCTATGCTCTGTATTTCCNAATGACAGGTCTTCCCAACACTTTGGTGAATGATGAA 1292  
QY 962 CGCTTCTATGATGGCCAAAGCACCATAAGGATTTTGTCTATGGATGGCTTGGTCAATATAG 1021  
DB 1293 CGCTTCTATGATGGCCAAAGCACCATAAGGATTTTGTCTATGNATGGCTTGGTCAATATAG 1352  
QY 1022 TTGGAGGATGCTGGGTCAACACAGATCATATCAAGGAAATTTCTGAAAGCTGTGAAA 1081  
DB 1353 TTGGAGGATGCTGGGTCAACACAGATCATATCAAGGAAATTTCTGAAAGCTGTGAAA 1412  
QY 1082 ATTGTAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGCTCGTC 1141  
DB 1413 ATTGTAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGCTCGTC 1472  
QY 1142 TAGAGCCTTTCAGGATTTGGACCGTACACCACTTTTGTAACTTTGGAGAGCGCTGTAAATG 1201  
DB 1473 TAGAGCCTTTCAGGATTTGGACCGTACACCACTTTTGTAACTTTGGAGAGCGCTGTAAATG 1532  
QY 1202 TTGCAGGATCAAGGAAGTTTGTCTAACTCATPATGCGCAGGAAATCATGAAAGACCTTGT 1261  
DB 1533 TTGCAGGATCAAGGAAGTTTGTCTAACTCATPATGCGCAGGAAATCATGAAAGACCTTGT 1592  
QY 1262 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCGAGGTGTGGATGTCAACATGATGATG 1321  
DB 1593 GTGTTGCCAAAGTGCAGTGGAAATGGAGCCGAGGTGTGGATGTCAACATGATGATG 1652  
QY 1322 GCATGCTAGATGTTCCAAAGTGCATGACCAATTTTGCATTTAACTTTTCCAGGACG 1381  
DB 1653 GCATGCTAGATGTTCCAAAGTGCATGACCAATTTTGCATTTAACTTTTCCAGGACG 1712  
QY 1382 ACATGCGAAAGTACCTTTTGTGCAATGCACTCCTCCAAATTTTGTGATGAAAGCTGGT 1441  
DB 1713 ACATGCGAAAGTACCTTTTGTGCAATGCACTCCTCCAAATTTTGTGATGAAAGCTGGT 1772  
QY 1442 TAAAGTCTGCGCAAGGAGTGCATTTGATGCAATGCAATTTAGTCTGAGGAGGAGGACG 1501  
DB 1773 TAAAGTCTGCGCAAGGAGTGCATTTGATGCAATGCAATTTAGTCTGAGGAGGAGGACG 1832  
QY 1502 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGATGATGAGTGTGCTATGGTGTGCTATGGCTT 1561  
DB 1833 ACTTCTTGGAGAGGCGCAGGAAGATTTAAAGATGATGAGTGTGCTATGGTGTGCTATGGCTT 1892  
QY 1562 TTGATGAAGAAGGACAGGACACAGAAACAGACACAAAAATCAGAGTGTGCAACCGGGCT 1621  
DB 1893 TTGATGAAGAAGGACAGGACACAGAAACAGACACAAAAATCAGAGTGTGCAACCGGGCT 1952  
QY 1622 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGCAATTTTTCAGCCTTAATA 1681  
DB 1953 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGCAATTTTTCAGCCTTAATA 2012  
QY 1682 TCCTAACCATTTGGGACTGGAATGGAGGAACCAACTGTATGCCAATTAATTTTATCCATG 1741  
DB 2013 TCCTAACCATTTGGGACTGGAATGGAGGAACCAACTGTATGCCAATTAATTTTATCCATG 2072  
QY 1742 CAACAAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACT 1801  
DB 2073 CAACAAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGAAATAGTGGAGGTCTTTTCCAACT 2132  
QY 1802 TGTCTTCTCTTCCAGGAATGGAGGCAATTCGAGAGCAATGATCGTGGGTCTTTCCTTT 1861

DB 2133 TGTCTTCTCTTCCAGGAATGGAGGCAATTCGAGAGCAATGATCGGTCTTTCCTTT 2192  
QY 1862 ACCATGCAATCAAGTCTGCGATGGAACATGAGAGATGATGAATGCTGAAACCTTCCCTGTGT 1921  
DB 2193 ACCATGCAATCAAGTCTGCGATGGAACATGAGGATGATGAATGCTGAAACCTTCCCTGTGT 2252  
QY 1922 ATGATGATATCCATAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTCGAATAAAGACC 1981  
DB 2253 ATGATGATATCCATAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTCGAATAAAGACC 2312  
QY 1982 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCAAGAGGAGGAGAAAG 2041  
DB 2313 CTGAGGCCACTGAGAAAGCTCTTACGTTATGCCCAGACTCAAGGCAAGAGGAGGAGAAAG 2372  
QY 2042 TCATTGAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACCTTGTAGTATGCCCCTTG 2101  
DB 2373 TCATTGAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACCTTGTAGTATGCCCCTTG 2432  
QY 2102 TGAAGGCGCATTTGAAAAACATATTTTGAAGTACTGAGGAGCCAGGTTAAACCAAAAA 2161  
DB 2433 TGAAGGCGCATTTGAAAAACATATTTTGAAGTACTGAGGAGCCAGGTTAAACCAAAAA 2492  
QY 2162 AATATCCCCGACCTCTCAATATATTTTGAAGGACCCCTGATGAATGGAATGAAAAATTTGTTG 2221  
DB 2493 AATATCCCCGACCTCTCAATATATTTTGAAGGACCCCTGATGAATGGAATGAAAAATTTGTTG 2552  
QY 2222 GTGATCTTTTGGAGCTGGAAAAATTTTCTACCTCAGGTTATAAAGTCAAGCCCGGGTTA 2281  
DB 2553 GTGATCTTTTGGAGCTGGAAAAATTTTCTACCTCAGGTTATAAAGTCAAGCCCGGGTTA 2612  
QY 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGAAAAAGAGAGAGAACCCAGAG 2341  
DB 2613 TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGAAAAAGAGAGAGAACCCAGAG 2672  
QY 2342 TGCTTAAAGCGCACAGTAGAAGAGAGAGAGCCCTTACAGGCGCACATCTGCTGGCCACTG 2401  
DB 2673 TGCTTAAAGCGCACAGTAGAAGAGAGAGAGCCCTTACAGGCGCACATCTGCTGGCCACTG 2732  
QY 2402 TTAAGGCGCAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTTGGCTGCAATA 2461  
DB 2733 TTAAGGCGCAGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTCTTGGCTGCAATA 2792  
QY 2462 ATTTCCGAGTTATTTAGGAGTCAATGATCCATGTGATAGTACTCAAGAGCTGCTC 2521  
DB 2793 ATTTCCGAGTTATTTAGGAGTCAATGATCCATGTGATAGTACTCAAGAGCTGCTC 2852  
QY 2522 TTGACCAAAAGCAGATATATTTGGCCTGTGAGGACTCATCACTCTCTTCCCTGGATGAAA 2581  
DB 2853 TTGACCAAAAGCAGATATATTTGGCCTGTGAGGACTCATCACTCTCTTCCCTGGATGAAA 2912  
QY 2582 TGATTTTGTTCAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATTTGGAG 2641  
DB 2913 TGATTTTGTTCAGGAAATGGAGAGATTAGCTATAAGGATTCATTTGTTGATTTGGAG 2972  
QY 2642 GAGCAACACTTCAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGACCTG 2701  
DB 2973 GAGCAACACTTCAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGACCTG 3032  
QY 2702 TAATCCATGTCTCGAGCGCTCCAAAGAGTGTGGTGTGTTCACAGCTGTTAGATGAAA 2761  
DB 3033 TAATCCATGTCTCGAGCGCTCCAAAGAGTGTGGTGTGTTCACAGCTGTTAGATGAAA 3092  
QY 2762 ATCTAAAGATGATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGACC 2881  
DB 3093 ATCTAAAGATGATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGGACC 3152  
QY 2882 ATTTAGTGTCTCAAGGAGAGGAGATCTTACCCTTTAAGTCAAGCCAGAAAAAGTGGTT 2881  
DB 3153 ATTTAGTGTCTCAAGGAGAGGAGATCTTACCCTTTAAGTCAAGCCAGAAAAANTTGGTT 3212  
QY 2882 TCCAAAATGGATTTGGCTGTCTGAAACCTCACCCAGTGAAGCCACCGTTTATTATTTGGGACCCAGG 2941

Db 3213 TCAGATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACGCTTTATTGGGACCCAGG 3272  
Qy 2942 TCTTTGAAGACTATGACCTGCAGAACTGGTGGACTACATTGACTGAAGCCTTTCTTTG 3001  
Db 3273 TCTTTGAAGACTATGACCTGCAGAACTGGTGGACTACATTGACTGAAGCCTTTCTTTG 3332  
Qy 3002 ATGCTGGCAGCTCCGGGCAAGTACCGBATCGAGGCTTCCCAAGATATTTAAACACA 3061  
Db 3333 ATGCTGGCAGCTCCGGGCAAGTACCGBATCGAGGCTTCCCAAGATATTTAAACACA 3392  
Qy 3062 AAACAGTAGGTGGAGAGGCGCAGGAAGGTCTACGATGATGCCACAATATGCTGAACAC 3121  
Db 3393 AAACAGTAGGTGGAGAGGCGCAGGAAGGTCTACGATGATGCCACAATATGCTGAACAC 3452  
Qy 3122 TGATTAGTCAAAAGAACTCCGGCCCGGGGTGTGGTTGGTCTTGGCCAGCAGAGTA 3181  
Db 3453 TGATTAGTCAAAAGAACTCCGGCCCGGGGTGTGGTTGGTCTTGGCCAGCAGAGTA 3512  
Qy 3182 TCCAGAGCAGATTCACCTGTACGCAGAGGCTGTGTCGCCAGGCTGCAGAGCCCATAG 3241  
Db 3513 TCCAGAGCAGATTCACCTGTACGCAGAGGCTGTGTCGCCAGGCTGCAGAGCCCATAG 3572  
Qy 3242 CCACCTTCTATGGGTAAAGCAACAGGCTGAGGAAGGACTCTGCCAGCAGCGAGCCATCT 3301  
Db 3573 CCACCTTCTATGGGTAAAGCAACAGGCTGAGGAAGGACTCTGCCAGCAGCGAGCCATCT 3632  
Qy 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCCTGGCATCCGCTGACTCTGGGCTGT 3361  
Db 3633 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTCCTGGCATCCGCTGACTCTGGGCTGT 3692  
Qy 3362 TTGCCGTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3421  
Db 3693 TTGCCGTGCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3752  
Qy 3422 ACTACAGCAGCATCATGCTCAAGCGCTGAGGAGGACTCTGCCAGCAGCGAGCCCTTTGCGAAG 3481  
Db 3753 ACTACAGCAGCATCATGCTCAAGCGCTGAGGAGGACTCTGCCAGCAGCGAGCCCTTTGCGAAG 3812  
Qy 3482 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541  
Db 3813 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3872  
Qy 3542 TCGGAGCTCGAGAGTTGCGGTACAAGGCGATCGCCCGGCTCTGGTACCCAGCC 3601  
Db 3873 TCGGAGCTCGGAGCTCGGTACAAAGGCGATCGCCCGGCTCTGGTACCCAGCC 3932  
Qy 3602 AGCCCGACACACCGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3661  
Db 3933 AGCCCGACACACCGAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3992  
Qy 3662 GCATTAGGTTAAGAGATCATTAGCAATGGCAGCTGCTTCCAGAGCTCTCAGGCTCTTACT 3721  
Db 3993 GCATTAGGTTAAGAGATCATTAGCAATGGCAGCTGCTTCCAGAGCTCTCAGGCTCTTACT 4052  
Qy 3722 TCTCCAAATTTGAAGTCAAAATATTTGCTGTGGGAGAAATTTCCAAAGGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGAAGTCAAAATATTTGCTGTGGGAGAAATTTCCAAAGGATCAGGTTGAGG 4112  
Qy 3782 ATTATGATTTGAGGAAGAAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAGCCATTT 3841  
Db 4113 ATTATGATTTGAGGAAGAAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAGCCATTT 4172  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3875  
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 4206

RESULT 16  
HSM807075 5240 bp mRNA linear PRI 17-APR-2005  
LOCUS Homo sapiens mRNA; cDNA DKFZp686A22123 (from clone DKFZp686A22123).  
DEFINITION BX641132  
ACCESSION BX641132.1 GI:34365460  
VERSION

## KEYWORDS

SOURCE  
ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

1 (bases 1 to 5240)

Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A.,  
Fobo, G., Han, M. and Wiemann, S.

The German Human cDNA Consortium

Direct Submission

Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANYClone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
consortium of the German Genome Project.

This clone (DKFZp686A22123) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at http://mips.gsf.de/proj/cDNA/.

## FEATURES

source

1..5240

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="RZPD:DKFZp686A22123Q"

/db\_xref="taxon:9606"

/map="1q43"

/clone="DKFZp686A22123"

/tissue\_type="human uterus"

/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_sfi; host

DH10B; sites SfiIA + SfiIB"

/dev\_stage="adult"

..5240

/gene="DKFZp686A22123"

Join(&lt;1..759,759..2459)

/gene="DKFZp686A22123"

/note="5-methyltetrahydrofolate--homocysteine

methyltransferase, N-terminus truncated

The frame shift was determined manually"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="CAE46059.1"

/db\_xref="GI:34365461"

/db\_xref="GOA:O6MZH4"

/db\_xref="InterPro:IPR000489"

/db\_xref="InterPro:IPR003759"

/db\_xref="InterPro:IPR004223"

/db\_xref="InterPro:IPR006158"

/db\_xref="InterPro:IPR011822"

/db\_xref="UniProt/TREMBL:Q6MZH4"

/translation="GMLDPSAMTRFCNLIASEPDIKVPCLIDSNPAVIEAGLKCC

QKQIVNSISLKGQATETDKIRVCTRAYHLVKKLGFNPNDIIPDNILITIGTME

ENLVNFIHATVTKETPGARISSGLNLSFSPRGMETREANHGVLVYHAIKSG

MDMGVFNAGLNPVYDIHKKELQLEDLIWNKDPKATEKLRYAQGTGGKKVIOTD

ENRNGVBERLEVAVKGIKHIIDTEARLNOKYPRPLNIEGLPMNGMKIYVDL

FCAGKMFLOVTKSARVKKAVGHILPMEKERESTVRLNGTVESDPYQGTIVLATV

KGDVHDIGKIVGVLCNNFRVIDLGVNTPCDKILKAALDKHADIIGLSGLITPSLD

EMIFVAKENERLAIRIPLIGGATSKTHTVKIAPRYSAPVHVLDAKSVVVCSQL

LDENLDEYFEEIMEEYIRQGHYESLKERRYLPQSARKSGFQMDLWPEHPVKPT

FTGTQVFDYDLQKLVYIDMKPFDVQMLRGKYPNRPKIFNDKTVGGEARKVYDD

AHNMLNTLISQKLRGVGVFWPAQSIQDDIHLVAAVPOAAEPIATFYGLRQOAE

KDSASTEPPYCLSDFTAPLHSGIRDYGLPAVACFGVSELSKAYEDDGDYSSIMWKA

LGRLAEPASSELHERVRELWAYCSEQLDVADLRRLYKIGIRPAPGPPSPQDHTKEK

LTMRLADIQSTGIRLTRELAWAPASVSGLYFNLKSKYFAVGKISKQDVEDYALR

KNISVAEVEKMLGPILGYDTD"

5208..5213

/gene="DKFZp686A22123"

5221

/gene="DKFZp686A22123"

polya\_signal

polya\_site

ORIGIN



Query Match 45.4%; Score 1779; DB 8; Length 5240; Best Local Similarity 99.6%; Pred. No. 0; Matches 2299; Conservative 0; Mismatches 8; Indels 1; Gaps 1;			
QY	1568	AAGAAGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCTACCATC	1627
DB	167	AAGAAGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCTACCATC	226
QY	1628	TGCTTGTGAAAAACTGGGCTTTAAATCCAAATGACATTTATTTTGACCCCTTAATCTCTAA	1687
DB	227	TGCTTGTGAAAAACTGGGCTTTAAATCCAAATGACATTTATTTTGACCCCTTAATCTCTAA	286
QY	1688	CCATTTGGACTGGAATGGAGGAAACAAACTTTGTATGCAATTAATTTTATCCATGCAACAA	1747
DB	287	CCATTTGGACTGGAATGGAGGAAACAAACTTTGTATGCAATTAATTTTATCCATGCAACAA	346
QY	1748	AAGTCATTAAGAACAATTAACCTGGAGCCAGAATAAGTGAGGTCTTTTCCAACTGTCTCT	1807
DB	347	AAGTCATTAAGAACAATTAACCTGGAGCCAGAATAAGTGAGGTCTTTTCCAACTGTCTCT	406
QY	1808	TCTCTCTCGAGGATGGAAGCCATTCGAGAAGCAATGCAATGGGTTTCTCTTTACCATG	1867
DB	407	TCTCTCTCGAGGATGGAAGCCATTCGAGAAGCAATGCAATGGGTTTCTCTTTACCATG	466
QY	1868	CAATCAAGTCTGGCATGGACATGAGATAGTGAATGCTGAAACCTCCCTGTGTATGATG	1927
DB	467	CAATCAAGTCTGGCATGGACATGAGATAGTGAATGCTGAAACCTCCCTGTGTATGATG	526
QY	1928	ATATCCATAAGGAACCTTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGG	1987
DB	527	ATATCCATAAGGAACCTTCGAGCTCTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGG	586
QY	1988	CCACTGAGAGCTCTTAAGTTATGCCAGACTCAAGGCAAGGAGGGAAGAAAGTCAATTC	2047
DB	587	CCACTGAGAGCTCTTAAGTTATGCCAGACTCAAGGCAAGGAGGGAAGAAAGTCAATTC	646
QY	2048	AGACTGATGAGTGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCTTGTGAAGG	2107
DB	647	AGACTGATGAGTGAGAAATGGCCCTGTGGAAGAACGCTTGAGTATGCCCTTGTGAAGG	706
QY	2108	GCATTTGAAAAACATTAATTTAGGATTAAGGAAAGCCAGGTAAACCAAAAAAATATC	2167
DB	707	GCATTTGAAAAACATTAATTTAGGATTAAGGAAAGCCAGGTAAACCAAAAAAATATC	765
QY	2168	CCCGACCTCTCAATTAATTTGAAGGACCCCTGTGATGAATGGAATGGAATTTGGTGATC	2227
DB	766	CCCGACCTCTCAATTAATTTGAAGGACCCCTGTGATGAATGGAATGGAATTTGGTGATC	825
QY	2228	TTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTATAAGTCAAGCCGCGGTTATGAAGA	2287
DB	826	TTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTATAAGTCAAGCCGCGGTTATGAAGA	885
QY	2288	AGGCTGTTGGGCACTTATCCCTTTCTGGAATAAGGAAAGAAACCAAGAGTGCTTA	2347
DB	886	AGGCTGTTGGGCACTTATCCCTTTCTGGAATAAGGAAAGAAACCAAGAGTGCTTA	945
QY	2348	ACGGCACAGTGAAGAAGAGGACCCCTTACAGGGCAACCATCGTGTGCGCCACTGTTAAAG	2407
DB	946	ACGGCACAGTGAAGAAGAGGACCCCTTACAGGGCAACCATCGTGTGCGCCACTGTTAAAG	1005
QY	2408	CGGAGCTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATATTTCC	2467
DB	1006	CGGAGCTGCACACATAGGCAAGAACATAGTTGGAGTAGTCTTGGCTGCAATATTTCC	1065
QY	2468	GAGTTATTGATTTAGGAGTCAATGACTCCATGTGATAAGATACTGAAAGCTGCTTTGACC	2527
DB	1066	GAGTTATTGATTTAGGAGTCAATGACTCCATGTGATAAGATACTGAAAGCTGCTTTGACC	1125
QY	2528	ACAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCATCTTTCCTGGAATGAATGATT	2587
DB	1126	ACAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCATCTTTCCTGGAATGAATGATT	1185
QY	2588	TTGTTGCCAAGGAATGGAGAGATTAGCTATAAGGATTTCCATTTGTTGATGGAGGACAA	2647
DB	1186	TTGTTGCCAAGGAATGGAGAGATTAGCTATAAGGATTTCCATTTGTTGATGGAGGACAA	1245
QY	2648	CCACTTTCAAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGCACCTGTAATCC	2707
DB	1246	CCACTTTCAAAAAACCCACACAGCAGTTTAAATAGCTCCGAGATACAGTGCACCTGTAATCC	1305
QY	2708	ATGTCCTGGAACGCGTCCAAAGAGTGTGGTGTGTTTCCAGCTGTGTAGATGAAATCTAA	2767
DB	1306	ATGTCCTGGAACGCGTCCAAAGAGTGTGGTGTGTTTCCAGCTGTGTAGATGAAATCTAA	1365
QY	2768	AGATGAAATACCTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGACCATATG	2827
DB	1366	AGATGAAATACCTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGACCATATG	1425
QY	2828	AGTCTCTCAAGGAGAGGAGATACCTTACCTTAAGTCAAGCCAGCAAAAAAGTGTTCCTAA	2887
DB	1426	AGTCTCTCAAGGAGAGGAGATACCTTACCTTAAGTCAAGCCAGCAAAAAAGTGTTCCTAA	1485
QY	2888	TGGATTGGCTGTCTGAAACCTCCACAGTGAAGCCCACTTATTTGGAACCCAGGTCTTTG	2947
DB	1486	TGGATTGGCTGTCTGAAACCTCCACAGTGAAGCCCACTTATTTGGAACCCAGGTCTTTG	1545
QY	2948	AAGACTATGACCTGTGCAAGAGCTGGTGGACTACATTTGACTGGAAGCCCTTTTGTGATGCT	3007
DB	1546	AAGACTATGACCTGTGCAAGAGCTGGTGGACTACATTTGACTGGAAGCCCTTTTGTGATGCT	1605
QY	3008	GGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTACGACAAACAG	3067
DB	1606	GGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTACGACAAACAG	1665
QY	3068	TAGTGTGAGAGGCGACGAAAGTCTACGATGATGCCCCCAATATGCTGGAACACACTGATTA	3127
DB	1666	TAGTGTGAGAGGCGACGAAAGTCTACGATGATGCCCCCAATATGCTGGAACACACTGATTA	1725
QY	3128	GTCAAAAGAAACTCCGGGCGCCGGGTGTGGTTTGGCTTCTGGCCAGCACAGAGTATCAAAG	3187
DB	1726	GTCAAAAGAAACTCCGGGCGCCGGGTGTGGTTTGGCTTCTGGCCAGCACAGAGTATCAAAG	1785
QY	3188	ACGACATTAACCTGTAGCGAGAGCTGTGTGCCCCCAGGCTGCAGAGCCCATAGCCACTT	3247
DB	1786	ACGACATTAACCTGTAGCGAGAGCTGTGTGCCCCCAGGCTGCAGAGCCCATAGCCACTT	1845
QY	3248	TCATATGGTTTAAAGCAACAGGCTGAGAGGACTCTGCGACGAGGACCATCTACTGCC	3307
DB	1846	TCATATGGTTTAAAGCAACAGGCTGAGAGGACTCTGCGACGAGGACCATCTACTGCC	1905
QY	3308	TCTCAGACTTCACTGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCGCTGTTGCCG	3367
DB	1906	TCTCAGACTTCACTGCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCGCTGTTGCCG	1965
QY	3368	TTGCTCTCTTTGGGTTAAGAGAGCTGAGCAAGGCGCTATGAGGATGATGGTGAAGCTACA	3427
DB	1966	TTGCTCTCTTTGGGTTAAGAGAGCTGAGCAAGGCGCTATGAGGATGATGGTGAAGCTACA	2025
QY	3428	GCAGCATCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCGCTTTGCAAGAGAGTCC	3487
DB	2026	GCAGCATCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCGCTTTGCAAGAGAGTCC	2085
QY	3488	ATGAAAGAGTTTCGCCGAGAACTGTGGGCGCTACTGTGGCAGTGAAGAGCTGAGCGTCCAG	3547
DB	2086	ATGAAAGAGTTTCGCCGAGAACTGTGGGCGCTACTGTGGCAGTGAAGAGCTGAGCGTCCAG	2145
QY	3548	ACCTGCGAAAGTTGCGGTACAAGGCGCATCCGCGCGGCTCTCTGGTACCCAGCCAGCGCG	3607
DB	2146	ACCTGCGAAAGTTGCGGTACAAGGCGCATCCGCGCGGCTCTCTGGTACCCAGCCAGCGCG	2205
QY	3608	ACCAACCGGAGAGCTCAACATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGCATTA	3667
DB	2206	ACCAACCGGAGAGCTCAACATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGCATTA	2265
QY	3668	GGTTTAAACAGAAATCAATTAGCAATGGCACCTGTGTTCCAGCAGTCTCTCAGGCGCTCTTCTCCA	3727



Db 2366 GGTTACAGAAATCATTAGCAATGSCACCTGCTTCAGCAGTCTCAGGCCTCTACTTCTCCA 2325  
Qy 3728 ATTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAGAGATCAGGTTGAGATTAATG 3787  
Db 2326 ATTGAAGTCCAAATATTTGCTGTGGGGAAGATTTCCAGAGATCAGGTTGAGATTAATG 2385  
Qy 3788 CATTGAGGAAGAACATATCTGTGCTCAGGTTGAGAAATGCTTTGGACCCATTTTGGAT 3847  
Db 2386 CATTGAGGAAGAACATATCTGTGCTCAGGTTGAGAAATGCTTTGGACCCATTTTGGAT 2445  
Qy 3848 ATGATACAGACTACTTTTTTTTTTTTTTTT 3875  
Db 2446 ATGATACAGACTACTTTTTTTTTTTTTTT 2473

RESULT 17  
CQ721809  
LOCUS 1835 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 7743 from Patent WO02068579.  
ACCESSION CQ721809  
VERSION CQ721809.1 GI:42282666  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
Patent: WO 02068579-A 7743 06-SEP-2002;  
PE Corporation (NY) (US)  
LOCATION/Qualifiers  
SOURCE 1..1835  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

FEATURES  
source  
ORIGIN  
Query Match 18.2%; Score 714; DB 6; Length 1835;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1252 GAAGCCTTGTGTGTTGCCAAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCAAC 1311  
Db 1071 GAAGCCTTGTGTGTTGCCAAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCAAC 1130

Qy 1312 ATGATGATGGATGCTTAGATGGTCCAAAGTGCATGACCAATGACCAATTTTGCAACTTAATTGCT 1371  
Db 1131 ATGATGATGGATGCTTAGATGGTCCAAAGTGCATGACCAATGACCAATTTTGCAACTTAATTGCT 1190

Qy 1372 TCCGAGCAGATCGCAAGGTACCTTTGTGATCGACTCTCCCAATTTTGTGTGATT 1431  
Db 1191 TCCGAGCAGATCGCAAGGTACCTTTGTGATCGACTCTCCCAATTTTGTGTGATT 1250

Qy 1432 GAAGCTGGGTTAAAGTCTGCAAGGAAGTGCATTGTCATAGCATTAGTCTGAGGAA 1491  
Db 1251 GAAGCTGGGTTAAAGTCTGCAAGGAAGTGCATTGTCATAGCATTAGTCTGAGGAA 1310

Qy 1492 GGAGAGGACGACTTTTGGAGAAAGGCCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTG 1551  
Db 1311 GGAGAGGACGACTTTTGGAGAAAGGCCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTG 1370

Qy 1552 GTCATGCTTTTGTATGAAGAGGACGAGCAACAGAAACAGACACAAAATCAGAGTGTGC 1611  
Db 1371 GTCATGCTTTTGTATGAAGAGGACGAGCAACAGAAACAGACACAAAATCAGAGTGTGC 1430

Qy 1612 ACCGGGSCCTACCATCTGCTTGTGAAAAAAGTGGCTTTTAAATCCAAATGACATTATTTT 1671  
Db 1431 ACCGGGSCCTACCATCTGCTTGTGAAAAAAGTGGCTTTTAAATCCAAATGACATTATTTT 1490

Qy 1672 GACCCTAATATCTTAACCATTTGGACTGGAATGGAGGAACAGAACTTGTATGCCATTAAAT 1731  
Db 1491 GACCCTAATATCTTAACCATTTGGACTGGAATGGAGGAACAGAACTTGTATGCCATTAAAT 1550  
Qy 1732 TTTATCCATGCAACAAAGTCATTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGGAGGT 1791  
Db 1551 TTTATCCATGCAACAAAGTCATTAAAGAAACATTTACCTGGAGCCAGAAATAAGTGGAGGT 1610  
Qy 1792 CTTTCCAACTGTCTCTCTTCCGAGGAATGGAAGCCATTCGAGAAGCAATATGCAATGGG 1851  
Db 1611 CTTTCCAACTGTCTCTCTTCCGAGGAATGGAAGCCATTCGAGAAGCAATATGCAATGGG 1670  
Qy 1852 GTTTTCTTTTACCATGCAATCAAGTCTGCGATGGAATGGAAGTGTGATGCTGGAAC 1911  
Db 1671 GTTTTCTTTTACCATGCAATCAAGTCTGCGATGGAATGGAAGTGTGATGCTGGAAC 1730  
Qy 1912 CTCCTCTGTATGATGATATCCATAAGGAATTTCTGAGCTCTGTGAAGATCTCATCTGG 1971  
Db 1731 CTCCTCTGTATGATGATATCCATAAGGAATTTCTGAGCTCTGTGAAGATCTCATCTGG 1790  
Qy 1972 AATAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGCCAG 2016  
Db 1791 AATAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGCCAG 1835

RESULT 18  
CQ850206  
LOCUS 4172 bp DNA linear PAT 23-AUG-2004  
DEFINITION Sequence 675 from Patent EP1447413.  
ACCESSION CQ850206  
VERSION CQ850206.1 GI:51508418  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
Isogai, T., Yamamoto, J., Nishikawa, T., Isono, Y., Sugiyama, T.,  
Otsuki, T., Wakamatsu, A., Ishii, S., Nagai, K. and Irie, R.  
Full-length human cDNA  
Patent: EP 1447413-A 675 18-AUG-2004;  
Research Association for Biotechnology (JP)  
LOCATION/Qualifiers  
SOURCE 1..4172  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 7.9%; Score 308; DB 6; Length 4172;  
Best Local Similarity 99.5%; Pred. No. 5.3e-168;  
Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3467 AGGCTTTTGAGAAGAGCTCCATGAAGAGTTGCGCAGAACTGTGGGCTACTGTGGCA 3526  
Db 1015 AGGCTTTTGAGAAGAGCTCCATGAAGAGTTGCGCAGAACTGTGGGCTACTGTGGCA 1074

Qy 3527 GTGAGCAGCTGCAAGCTGCGAGACCTTGGGTTACGAGGATCCGCGGCTC 3586  
Db 1075 GTGAGCAGCTGCAAGCTGCGAGACCTTGGGTTACGAGGATCCGCGGCTC 1134

Qy 3587 CTGGCTTACCCAGCAGCCGACCAACCGAGAAGCTCACCATGTGGAGATCGCAGACA 3646  
Db 1135 CTGGCTTACCCAGCAGCCGACCAACCGAGAAGCTCACCATGTGGAGATCGCAGACA 1194

Qy 3647 TCGAGCAGTCTACAGCATTAGGTTAAACAGAAATCATTAAGCAATGGCACTGTGAGCAG 3706  
Db 1195 TCGAGCAGTCTACAGCATTAGGTTAAACAGAAATCATTAAGCAATGGCACTGTGAGCAG 1254

Qy 3707 TCTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTCCA 3766  
Db 1255 TCTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTCCA 1314

QY 3767 AGGATCAGGTTGAGGATTATGCAATTGAGGAGACATATCTGTGGCTGAGGTTGAGAAAT 3826  
 Db 1315 AGGATCAGGTTGAGGATTATGCAATTGAGGAGACATATCTGTGGCTGAGGTTGAGAAAT 1374  
 QY 3827 GGCTTGAGCCCAATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3876  
 Db 1375 GGCTTGAGCCCAATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 1424

RESULT 19  
 AK127317 4172 bp mRNA linear PRI 19-FEB-2004  
 LOCUS Homo sapiens cDNA FLJ45386 fig, clone BRHIP3023922, highly similar  
 DEFINITION to 5-methyltetrahydrofolate--homocysteine methyltransferase (BC  
 2.1.1.13).

ACCESSION AK127317  
 VERSION AK127317.1 GI:34534177  
 KEYWORDS oligo capping; fig (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1  
 Kaneshiro, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,  
 Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,  
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
 Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,  
 Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
 Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,  
 Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.  
 and Isogai, T.

TITLE NEDO human cDNA sequencing project  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 4172)  
 AUTHORS Isogai, T. and Yamamoto, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
 (E-mail: genomics@hri.co.jp; Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: HRI and  
 RAB; annotation: Reverse Proteomics Research Institute, HRI and  
 RAB.

FEATURES  
 source Location/Qualifiers  
 1..4172  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="BRHIP3023922"  
 /tissue\_type="hippocampus"  
 /clone\_lib="BRHIP3"  
 /note="cloning vector: pME18SFL3"

ORIGIN  
 Query Match 7.9%; Score 308; DB 8; Length 4172;  
 Best Local Similarity 99.5%; Pred. No. 5.3e-168;  
 Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3467 AGGCTTTGCAAGAGCTCCATGAAGAGTTGCGCGGAACTGTGGGCTTACTGTGGCA 3526  
 Db 1015 AGGCTTTGCAAGAGCTCCATGAAGAGTTGCGCGGAACTGTGGGCTTACTGTGGCA 1074  
 QY 3527 GTGAGCAGCTGGACGCTCCAGACCTCGGAAGGTTGCGGTACAAGGGCATCCGCGGCTC 3586  
 Db 1075 GTGAGCAGCTGGACGCTCCAGACCTCGGAAGGTTGCGGTACAAGGGCATCCGCGGCTC 1134

QY 3587 CTGGCTACCCGAGCCGCGGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACA 3646  
 Db 1135 CTGGCTACCCGAGCCGCGGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACA 1194  
 QY 3647 TCGAGCAGTCTACAGGCATTAGGTTAAAGAAATCATTAGCAATGGCACCTGCTTCAGCAG 3706  
 Db 1195 TCGAGCAGTCTACAGGCATTAGGTTAAAGAAATCATTAGCAATGGCACCTGCTTCAGCAG 1254  
 QY 3707 TCTCAGGCTCTACTTCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTCCA 3766  
 Db 1255 TCTCAGGCTCTACTTCTCCAAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTCCA 1314  
 QY 3767 AGCATCAGGTTGAGGATTATGCAATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAAT 3826  
 Db 1315 AGCATCAGGTTGAGGATTATGCAATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAAT 1374  
 QY 3827 GGCTTGAGCCCAATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3876  
 Db 1375 GGCTTGAGCCCAATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTT 1424

RESULT 20  
 AY429553 435 bp mRNA linear PRI 27-APR-2004  
 LOCUS Homo sapiens MTRV\_1 mRNA sequence; alternatively spliced.  
 DEFINITION  
 ACCESSION AY429553  
 VERSION AY429553.1 GI:46488815  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 435)  
 Johnson, J.M., Castle, J., Garrett-Engle, P.W., Kan, Z., Loerch, P.M.,  
 Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and  
 Shoemaker, D.D.

TITLE Genome-wide survey of human alternative pre-mRNA splicing with exon  
 junction microarrays  
 JOURNAL Science 302 (5653), 2141-2144 (2003)  
 PUBMED 14684825  
 REFERENCE 2 (bases 1 to 435)  
 Johnson, J.M., Castle, J., Garrett-Engle, P.W., Loerch, P.M., Kan, Z.,  
 Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and  
 Shoemaker, D.D.

TITLE Direct Submission  
 JOURNAL Submitted (07-OCT-2003) Genomic Discovery, Rosetta/MRL, 12040 115th  
 Ave NE, Kirkland, WA 98034, USA

FEATURES  
 Location/Qualifiers  
 1..435  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 <1..>435  
 /note="MTRV\_1; alternatively spliced compared to  
 NM\_000254"

ORIGIN  
 Query Match 7.6%; Score 297; DB 8; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-161;  
 Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 530 CCCCATCTGTGGAAGCCCGGATATAGGAACATCATCTTGAGCTTGTGAAGCAT 589  
 Db 1 CCCCATCTGTGGAAGCCCGGATATAGGAACATCATCTTGAGCTTGTGAAGCAT 60  
 QY 590 ACCAAGAGAGCGCCAAAGGACTCTTCGATGCGGGGTTGATATCTTACTCATTAAGACTA 649  
 Db 61 ACCAAGAGAGCGCCAAAGGACTCTTCGATGCGGGGTTGATATCTTACTCATTAAGACTA 120  
 QY 650 TTTTGTGATCTGCGCAATGCCAAGCAGCCTTGTGTCATCCAAATCTTTTTCAGGAGA 709  
 Db 121 TTTTGTGATCTGCGCAATGCCAAGCAGCCTTGTGTCATCCAAATCTTTTTCAGGAGA 180

Qy 710 AATATGCTCCCGCCTTCTTTATTTTCAGGACGATCGTTGATATAAAAGTGGCGGACTC 769  
 Db 181 AATATGCTCCCGCCTTCTTTATTTTCAGGACGATCGTTGATATAAAAGTGGCGGACTC 240  
 Qy 770 TTTCGGACACAGAGGAGGATTTCTCATCCGCTCTCATGGAGAACCACTCT 826  
 Db 241 TTTCGGACACAGGAGGAGGATTTCTCATCCGCTCTCATGGAGAACCACTCT 297

RESULT 21  
 BD060282  
 LOCUS 305 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Secreted expressed sequence tags (SESTs).  
 ACCESSION BD060282  
 VERSION BD060282.1 GI:22605888  
 KEYWORDS JP 2001518793-A/642.  
 SOURCE Zea mays  
 ORGANISM Zea mays

REFERENCE 1 (bases 1 to 305)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Jacobs, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D., Treacy, M., Spaulding, V. and Agostino, M.J.  
 TITLE Secreted expressed sequence tags (SESTs)  
 JOURNAL Patent: JP 2001518793-A 642 16-OCT-2001;  
 GENETICS INSTITUTE INC  
 COMMENT PN JP 2001518793-A/642  
 PD 16-OCT-2001  
 PF 10-APR-1998 JP 1998543070  
 PR 10-APR-1997 US 08/837312  
 PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI DAVID MERBERG,  
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
 C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers

FEATURES  
 source 1..305  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4577"

ORIGIN  
 Query Match 5.7%; Score 222; DB 6; Length 305;  
 Best Local Similarity 99.6%; Pred. No. 2e-117;  
 Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2751 GTTAGATGAAATCTAAGGATGAATCTTTGAGGAATCATGGAAGATATGAAGATAT 2810  
 Db 22 GTTAGATGAAATCTAAGGATGAATCTTTGAGGAATCATGGAAGATATGAAGATAT 81

Qy 2811 TAGACAGACCATATGAGTCTCTCAAGGAGAGAGATATCTTACCTTAAGTCAAGCCAG 2870  
 Db 82 TAGACAGGCGCATATGAGTCTCTCAAGGAGAGAGATATCTTACCTTAAGTCAAGCCAG 141

Qy 2871 AAAAAGTGGTTTCAAAATGATGGCTGTCTGACCTCCACCGATGAAGCCAGTTTAT 2930  
 Db 142 AAAAAGTGGTTTCAAAATGATGGCTGTCTGACCTCCACCGATGAAGCCAGTTTAT 201

Qy 2931 TGGGACCCAGTCTTTGAAGACTATGACCTGCAGAGCTGTGTGACTACATGACTGGAA 2990  
 Db 202 TGGGACCCAGTCTTTGAAGACTATGACCTGCAGAGCTGTGTGACTACATGACTGGAA 261

Qy 2991 GCCTTTCTTTGATCTCTGGCAGCTCCGGGGCAA 3023  
 Db 262 GCCTTTCTTTGATCTCTGGCAGCTCCGGGGCAA 294

RESULT 22  
 AL359185

LOCUS AL359185 108841 bp DNA linear PRI 18-MAY-2005  
 DEFINITION Human DNA sequence from clone RP11-382M16 on chromosome 1. Contains the 3' end of the ACTN2 gene for actinin, alpha 2, the 5' end of the MTR gene for 5-methyltetrahydrofolate-homocysteine methyltransferase, a laminin receptor 1 (ribosomal protein SA, 67kDa) (LAMR1) pseudogene and a CpG island, complete sequence.

ACCESSION AL359185  
 VERSION AL359185.25 GI:18855190  
 KEYWORDS HTG; ACTN3; CpG island; LAMR1; MTR.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 108841)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 Skuce, C.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

COMMENT Clone requests: clonerequest@sanger.ac.uk  
 On Feb 21, 2002 this sequence version replaced gi:18655978.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chrl  
 RP11-382M16 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACE3.6

----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: vegas@sanger.ac.uk  
 -----

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES  
 source Location/Qualifiers  
 1..108841  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="1"  
 /clone="RP11-382M16"  
 /clone\_lib="RPC1-11.2"  
 join(AL359921.13:163608..163899,  
 AL359921.13:194958..195072,AL359921.13:195994..196113,  
 AL359921.13:197205..197291,AL359921.13:203033..203120,  
 AL359921.13:204778..204856,AL359921.13:208333..208414,  
 977..1062,2464..2556,4644..4874,8238..8385,9968..10118,  
 13009..13117,14466..14606,16812..16994,19289..19423,  
 20361..20540,22828..22974,25066..25131,26357..26515,  
 27803..29973)  
 /gene="ACTN2"  
 /locus\_tag="RP11-382M16.1-001"  
 join(AL359921.13:163608..163899,  
 AL359921.13:194958..195072,AL359921.13:195994..196113,  
 AL359921.13:197205..197291,AL359921.13:203033..203120,  
 AL359921.13:204778..204856,AL359921.13:208333..208414,  
 977..1062,2464..2556,4644..4874,8238..8385,9968..10118,  
 13009..13117,14466..14606,16812..16994,19289..19423,  
 20361..20540,22828..22974,25066..25131,26357..26515,  
 27803..29973)

gene

mRNA





```

/misc_feature      /note="assembly_fragment:01320"
148913 164368
/note="assembly_fragment:01509"
164469 .167514
/note="assembly_fragment:01886"
167615 .179223
/note="assembly_fragment:01970"
179324 .183166
/note="assembly_fragment:02051"

ORIGIN
Query Match      5.5% Score 217; DB 14; Length 183166;
Best Local Similarity 100.0%; Pred. No. 1.7e-114;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 GAAGGTCGTAAGAAACCCCTGGCGGATGAGATCAATGCCATTCCTGCAGAGAGATTATG 156
|
Db 97547 GAAGGTCGTAAGAAACCCCTGGCGGATGAGATCAATGCCATTCCTGCAGAGAGATTATG 97606
|
Qy 157 GTGCTGATGAGGAGTGGGAGCCATCATCCAGCGGAGAGCTTAAAGAACACCTTC 216
|
Db 97607 GTGCTGATGAGGAGTGGGAGCCATCATCCAGCGGAGAGCTTAAAGAACACCTTC 97666
|
Qy 217 CGAGGTCAGGAATTTAAGATCATGCCAGGCGCTGAAAGGCAACAATGACATTTTAAGT 276
|
Db 97667 CGAGGTCAGGAATTTAAGATCATGCCAGGCGCTGAAAGGCAACAATGACATTTTAAGT 97726
|
Qy 277 ATAAGTCAGGCTGATGTCATTTACCAATCCATAAGG 313
|
Db 97727 ATAAGTCAGGCTGATGTCATTTACCAATCCATAAGG 97763

RESULT 24
AC079975/c
AC079975 164529 bp DNA linear HTG 22-NOV-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-11N21 map 1, WORKING DRAFT
SEQUENCE, 37 unordered pieces.
AC079975
AC079975.2 GI:11276222
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 164529)
Homo sapiens chromosome 1, clone RP11-11N21
Unpublished
2 (bases 1 to 164529)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,
Bouckgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirell,A., Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-SEP-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

## COMMENT

```

On Nov 22, 2000 this sequence version replaced gi:10198369.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11095
Center clone name: L1_N_21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 147566 bases at least Q40
Consensus quality: 155962 bases at least Q30
Consensus quality: 159155 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 160929; sum-of-contigs
Quality coverage: 3.7 in Q20 bases; agarose-fp
Quality coverage: 3.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 638: contig of 638 bp in length
* 639
* 738: gap of 100 bp
* 739
* 1280: contig of 542 bp in length
* 1281
* 1380: gap of 100 bp
* 1381
* 2021: contig of 641 bp in length
* 2022
* 2121: gap of 100 bp
* 2122
* 2830: contig of 709 bp in length
* 2831
* 2930: gap of 100 bp
* 2931
* 3970: contig of 1040 bp in length
* 3971
* 4070: gap of 100 bp
* 4071
* 4923: contig of 853 bp in length
* 4924
* 5023: gap of 100 bp
* 5024
* 6063: contig of 1040 bp in length
* 6064
* 6163: gap of 100 bp
* 6164
* 7446: contig of 1283 bp in length
* 7447
* 7546: gap of 100 bp
* 7547
* 9381: contig of 1835 bp in length
* 9382
* 9481: gap of 100 bp
* 9482
* 11219: contig of 1738 bp in length
* 11220
* 11219: gap of 100 bp
* 11220
* 12350: contig of 1031 bp in length
* 12351
* 12450: gap of 100 bp
* 12451
* 13291: contig of 841 bp in length
* 13292
* 13391: gap of 100 bp
* 13392
* 14939: contig of 1548 bp in length
* 15039: gap of 100 bp
* 15040
* 17136: contig of 2097 bp in length
* 17137
* 17236: gap of 100 bp
* 17237
* 19620: contig of 2384 bp in length
* 19621
* 19720: gap of 100 bp
* 19721
* 22736: contig of 3016 bp in length
* 22737
* 22836: gap of 100 bp
* 22837
* 45880: contig of 22944 bp in length
* 45881
* 45894: gap of 100 bp
* 45895
* 48594: contig of 2714 bp in length
* 48595
* 48694: gap of 100 bp
* 48695
* 51084: contig of 2390 bp in length
* 51085
* 51184: gap of 100 bp
* 51185
* 53490: contig of 2306 bp in length
* 53491
* 53590: gap of 100 bp
* 53591
* 55545: contig of 1955 bp in length

```



```
* 55546 55645: gap of 100 bp
* 55646 57961: contig of 2316 bp in length
* 57962 58061: gap of 100 bp
* 58062 59375: contig of 1314 bp in length
* 59376 59475: gap of 100 bp
* 59476 63078: contig of 3603 bp in length
* 63079 63178: gap of 100 bp
* 63179 65356: contig of 2178 bp in length
* 65357 65456: gap of 100 bp
* 65457 68583: contig of 3127 bp in length
* 68584 68683: gap of 100 bp
* 73074 73174: contig of 4391 bp in length
* 73175 76775: contig of 3501 bp in length
* 76776 76775: gap of 100 bp
* 76776 81036: contig of 4261 bp in length
* 81037 81136: gap of 100 bp
* 81137 86818: contig of 5682 bp in length
* 86819 86918: gap of 100 bp
* 86919 94761: contig of 7843 bp in length
* 94762 94861: gap of 100 bp
* 94862 102807: contig of 7945 bp in length
* 102807 102906: gap of 100 bp
* 102907 117954: contig of 15048 bp in length
* 117955 118054: gap of 100 bp
* 118055 130927: contig of 12873 bp in length
* 130928 131028: gap of 100 bp
* 131028 145686: contig of 14659 bp in length
* 145687 145786: gap of 100 bp
* 145787 162777: contig of 16991 bp in length
* 162778 162878: gap of 100 bp
* 162878 164529: contig of 1652 bp in length.
```

## FEATURES

## Source

```
1. .164529
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="1"
/clone="RP11-11N21"
/clone_lib="RP11-11 Human Male BAC"
1. .638
/notes="assembly_fragment"
clone_end:SP6
vector_side:left
639. .738
/estimated_length=100
739. .1280
/notes="assembly_fragment"
1281. .1380
/estimated_length=100
1381. .2021
/notes="assembly_fragment"
2022. .2121
/estimated_length=100
2122. .2830
/notes="assembly_fragment"
2831. .2930
/estimated_length=100
2931. .3970
/notes="assembly_fragment"
3971. .4070
/estimated_length=100
4071. .4923
/notes="assembly_fragment"
4924. .5023
/estimated_length=100
5024. .6063
/notes="assembly_fragment"
6064. .6163
/estimated_length=100
6164. .7446
/notes="assembly_fragment"
```

```
gap 7447. .7546
/estimated_length=100
misc_feature 7547. .9381
/notes="assembly_fragment"
gap 9382. .9481
/estimated_length=100
misc_feature 9482. .11219
/notes="assembly_fragment"
gap 11220. .11319
/estimated_length=100
misc_feature 11320. .12350
/notes="assembly_fragment"
gap 12351. .12450
/estimated_length=100
misc_feature 12451. .13291

Query Match 5.2%; Score 204; DB 14; Length 164529;
Best Local Similarity 100.0%; Pred. No. 7.6e-107;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3266 AGGCTGAGAGGACTCTGCCAGCAGGAGCCATACTACTGCTCTCAGACTTCATCGCTC 3325
DB 41412 AGGCTGAGAGGACTCTGCCAGCAGGAGCCATACTACTGCTCTCAGACTTCATCGCTC 41353
QY 3326 CCTTGCATTCTGGCATCCGTGACTACCTGGGCGCTGTTGGCCGTTGCTGCTTTGGGGTAG 3385
DB 41352 CCTTGCATTCTGGCATCCGTGACTACCTGGGCGCTGTTGGCCGTTGCTGCTTTGGGGTAG 41293
QY 3386 AAGAGCTGAGCAAGCGCTATGAGGATGATGGTGAGCAGCTACAGCAGCATCGTCAAGG 3445
DB 41292 AAGAGCTGAGCAAGCGCTATGAGGATGATGGTGAGCAGCTACAGCAGCATCGTCAAGG 41233
QY 3446 CGCTGGGGACCGGCTGGCAGAGG 3469
DB 41232 CGCTGGGGACCGGCTGGCAGAGG 41209
```

## RESULT 25

## AC074231

```
LOCUS AC074231 164933 bp DNA linear HTG 21-JUL-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-192B19, WORKING DRAFT
SEQUENCE, 9 ordered pieces.
```

## ACCESSION AC074231

VERSION AC074231.1 GI:9309505

HTG; HTGS PHASE2; HTGS\_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 164933)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 1

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 164933)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (21-JUL-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

## COMMENT

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

-----

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.





QY 3446 CGCTGGGGGACCGCTGGCAGAGG 3469  
 Db 53036 CGCTGGGGGACCGCTGGCAGAGG 53059

RESULT 27  
 LOCUS AC080149/c  
 DEFINITION Homo sapiens chromosome 14 clone RP11-10110 map 14, WORKING DRAFT  
 SEQUENCE, 21 unordered pieces.  
 AC080149 GI:11276221  
 VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 176671)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 14, clone RP11-10110  
 Unpublished  
 2 (bases 1 to 176671)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,  
 Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
 Choepel,Y., Collangeo,M., Collins,S., Collymore,A., Cooke,P.,  
 DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
 FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
 Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
 Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
 Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
 O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,  
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
 Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
 Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strausz,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
 Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (26-SEP-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 176671)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,  
 Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
 Choepel,Y., Collangeo,M., Collins,S., Collymore,A., Cooke,P.,  
 DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
 FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
 Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
 Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
 Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
 O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,  
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
 Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
 Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strausz,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
 Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Nov 22, 2000 this sequence version replaced gi:10305243.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L11093  
 Center clone name: 10\_I\_10  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 169296 bases at least Q40  
 Consensus quality: 172583 bases at least Q30  
 Consensus quality: 173740 bases at least Q20  
 Insert size: 186000; agarose-fp  
 Insert size: 174671; sum-of-contigs  
 Quality coverage: 5.3 in Q20 bases; agarose-fp  
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 21 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 171: contig of 171 bp in length  
 \* 172 271: gap of 100 bp  
 \* 272 866: contig of 595 bp in length  
 \* 867 966: gap of 100 bp  
 \* 967 1687: contig of 701 bp in length  
 \* 1688 1767: gap of 100 bp  
 \* 1768 2728: contig of 961 bp in length  
 \* 2729 2828: gap of 100 bp  
 \* 2829 3495: contig of 667 bp in length  
 \* 3496 3595: gap of 100 bp  
 \* 3596 4294: contig of 699 bp in length  
 \* 4295 4394: gap of 100 bp  
 \* 4395 5091: contig of 697 bp in length  
 \* 5092 5191: gap of 100 bp  
 \* 5192 5814: contig of 623 bp in length  
 \* 5815 5914: gap of 100 bp  
 \* 5915 6561: contig of 646 bp in length  
 \* 6561 6660: gap of 100 bp  
 \* 6661 7288: contig of 628 bp in length  
 \* 7289 8033: contig of 645 bp in length  
 \* 8034 8133: gap of 100 bp  
 \* 8134 11132: contig of 3019 bp in length  
 \* 11153 11252: gap of 100 bp  
 \* 11253 16926: contig of 5674 bp in length  
 \* 16927 17026: gap of 100 bp  
 \* 17027 48247: contig of 31221 bp in length  
 \* 48248 48347: gap of 100 bp  
 \* 48348 56048: contig of 7701 bp in length  
 \* 56049 56149: gap of 100 bp  
 \* 56149 68743: contig of 12594 bp in length  
 \* 68743 68842: gap of 100 bp  
 \* 68843 86438: contig of 17596 bp in length  
 \* 86439 86538: gap of 100 bp  
 \* 86539 107928: contig of 21389 bp in length  
 \* 107928 108027: gap of 100 bp  
 \* 108028 135702: contig of 27675 bp in length  
 \* 135703 135802: gap of 100 bp  
 \* 135803 170129: contig of 34327 bp in length  
 \* 170130 170225: gap of 100 bp  
 \* 170226 176671: contig of 6442 bp in length.  
 Location/Qualifiers  
 1. 176671  
 /organism="Homo sapiens"

FEATURES  
 source



[illegible]



Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,  
Zimmer, A. and Zody, M.

# TITLE JOURNAL

REFERENCE  
AUTHORS

## Direct Submission

Submitted (26-SEP-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 176671)

Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L.,  
Boukhalil, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,  
Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferreira, P.,  
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,  
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,  
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRoque, K.,  
Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,  
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,  
McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,  
Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T.,  
O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,  
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,  
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,  
Zimmer, A. and Zody, M.

## TITLE JOURNAL

COMMENT

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 22, 2000 this sequence version replaced gi:10305243.  
All repeats were identified using RepeatMasker:  
Snit, A. F. A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L11093

Center Clone name: 10\_110

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 169296 bases at least Q40

Consensus quality: 172583 bases at least Q30

Consensus quality: 173740 bases at least Q20

Insert size: 186000; agarose-fp

Insert size: 174671; sum-of-contigs

Quality coverage: 5.3 in Q20 bases; agarose-fp

Quality coverage: 5.7 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 171: contig of 171 bp in length  
\* 172  
\* 172: gap of 100 bp  
\* 272  
\* 272: contig of 595 bp in length  
\* 867  
\* 867: gap of 100 bp  
\* 1667  
\* 1667: contig of 701 bp in length  
\* 1668  
\* 1668: gap of 100 bp  
\* 1768  
\* 1768: contig of 961 bp in length  
\* 2729  
\* 2729: gap of 100 bp  
\* 3496  
\* 3496: contig of 667 bp in length  
\* 3496

FEATURES	source	Location/Qualifiers
		1..176671
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosomes="14"
		/map="14"
		/clone="RP11-10110"
		/clone_lib="RPC1-11 Human Male BAC"
		1..171
		/note="assembly_fragment"
		clone_end:SP6
		vector_side:left
		172..271
		/estimated_length=100
		272..866
		/note="assembly_fragment"
		867..966
		/estimated_length=100
		967..1667
		/note="assembly_fragment"
		1668..1767
		/estimated_length=100
		1768..2728
		/note="assembly_fragment"
		2729..2828
		/estimated_length=100
		2829..3495
		/note="assembly_fragment"
		3496..3595
		/estimated_length=100
		3596..4294
		/note="assembly_fragment"
		4295..4394
		/estimated_length=100
		4395..5091
		/note="assembly_fragment"
		5092..5191
		/estimated_length=100
		5192..5814
		/note="assembly_fragment"
		5815..5914
		4294: contig of 699 bp in length
		4394: gap of 100 bp
		5091: contig of 697 bp in length
		5191: gap of 100 bp
		5814: contig of 623 bp in length
		5914: gap of 100 bp
		6560: contig of 646 bp in length
		6660: gap of 100 bp
		7288: contig of 628 bp in length
		7388: gap of 100 bp
		8033: contig of 645 bp in length
		8133: gap of 100 bp
		11152: contig of 3019 bp in length
		11252: gap of 100 bp
		16296: contig of 5674 bp in length
		17026: gap of 100 bp
		48247: contig of 31221 bp in length
		48347: gap of 100 bp
		56048: contig of 7701 bp in length
		56148: gap of 100 bp
		68742: contig of 12594 bp in length
		68842: gap of 100 bp
		86438: contig of 17596 bp in length
		86439: gap of 100 bp
		107927: contig of 21389 bp in length
		108027: gap of 100 bp
		135702: contig of 27675 bp in length
		135703: gap of 100 bp
		170129: contig of 34327 bp in length
		170130: gap of 100 bp
		176671: contig of 6442 bp in length.

```

misc_feature      /estimated_length=100
5915..6560
/note="assembly_fragment"
gap              6561..6660
/estimated_length=100
misc_feature      6661..7288
/note="assembly_fragment"
gap              7289..7388
/estimated_length=100
misc_feature      7389..8033
/note="assembly_fragment"
gap              8034..8133
/estimated_length=100
misc_feature      8134..11152
/note="assembly_fragment"
gap              11153..11252
/estimated_length=100
misc_feature      11253..16926
/note="assembly_fragment"
gap              16927..17026
/estimated_length=100
misc_feature      17027..48247

Query Match      4.7% Score 183; DB 14; Length 176671;
Best Local Similarity 100.0%; Pred. No. 1.7e-94;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1577 AGGCAACAGAAACAGACACAAAATCAGAGTGTGCACCCGGCGCTACCATCTGCTGTGA 1636
Db 76700 AGGCAACAGAAACAGACACAAAATCAGAGTGTGCACCCGGCGCTACCATCTGCTGTGA 76759

Qy 1637 AAAAAGTGGGCTTTAAATCCAAATGACATTTTGTGACCCCTATATCTTAACCATTTGGA 1696
Db 76760 AAAAAGTGGGCTTTAAATCCAAATGACATTTTGTGACCCCTATATCTTAACCATTTGGA 76819

Qy 1697 CTGGAATGGAGGACACAACTTGATGCACTTAATTTTATCCATGCAACAAAGTCATTA 1756
Db 76820 CTGGAATGGAGGACACAACTTGATGCACTTAATTTTATCCATGCAACAAAGTCATTA 76879

Qy 1757 AAG 1759
Db 76880 AAG 76882

RESULT 30
AC016512/c
LOCUS              73656 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-11N21, LOW-PASS SEQUENCE SAMPLING.
ACCESSION          AC016512
VERSION            AC016512.3 GI:9123838
KEYWORDS            HTG; HTGS_PHASE0.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 73656)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguelavkiy,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLoughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

```

# TITLE JOURNAL COMMENT

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (01-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 13, 2000 this sequence version replaced gi:7144825.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3233  
Center clone name: 11\_N\_21  
-----

\* NOTE: This record contains 81 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1 809: contig of 809 bp in length  
\* 909: gap of 100 bp  
\* 910 1718: contig of 809 bp in length  
\* 1719 1818: gap of 100 bp  
\* 1819 2635: contig of 817 bp in length  
\* 2636 2735: gap of 100 bp  
\* 2736 3541: contig of 806 bp in length  
\* 3542 4430: contig of 789 bp in length  
\* 4431 4530: gap of 100 bp  
\* 4531 5317: contig of 787 bp in length  
\* 5318 5417: gap of 100 bp  
\* 5418 6230: contig of 813 bp in length  
\* 6231 7135: contig of 805 bp in length  
\* 7136 7235: gap of 100 bp  
\* 7236 8055: contig of 820 bp in length  
\* 8056 8155: gap of 100 bp  
\* 8156 8959: contig of 804 bp in length  
\* 8960 9059: gap of 100 bp  
\* 9060 9873: contig of 814 bp in length  
\* 9874 9973: gap of 100 bp  
\* 9974 10777: contig of 804 bp in length  
\* 10778 10877: gap of 100 bp  
\* 10878 11696: contig of 819 bp in length  
\* 11697 11796: gap of 100 bp  
\* 11797 12593: contig of 797 bp in length  
\* 12594 12693: gap of 100 bp  
\* 12694 13490: contig of 797 bp in length  
\* 13491 13590: gap of 100 bp  
\* 13591 14402: contig of 812 bp in length  
\* 14403 14502: gap of 100 bp  
\* 14503 15304: contig of 802 bp in length  
\* 15305 15404: gap of 100 bp  
\* 15405 16210: contig of 806 bp in length  
\* 16211 17114: contig of 804 bp in length  
\* 17115 17214: gap of 100 bp  
\* 17215 18042: contig of 828 bp in length  
\* 18043 18142: gap of 100 bp  
\* 18143 18958: contig of 816 bp in length  
\* 18959 19059: gap of 100 bp  
\* 19059 19870: contig of 812 bp in length  
\* 19871 19871: gap of 100 bp

```
* 19971 20775: contig of 805 bp in length
* 20776 20875: gap of 100 bp
* 20876 21691: contig of 816 bp in length
* 21692 21791: gap of 100 bp
* 21792 22615: contig of 824 bp in length
* 22616 22715: gap of 100 bp
* 22716 23553: contig of 837 bp in length
* 23553 23652: gap of 100 bp
* 23653 24462: contig of 810 bp in length
* 24463 24562: gap of 100 bp
* 24563 25381: contig of 819 bp in length
* 25382 25481: gap of 100 bp
* 25482 26299: contig of 818 bp in length
* 26300 26399: gap of 100 bp
* 26400 27217: contig of 818 bp in length
* 27218 27317: gap of 100 bp
* 27318 28122: contig of 805 bp in length
* 28123 28222: gap of 100 bp
* 28223 29038: contig of 816 bp in length
* 29039 29138: gap of 100 bp
* 29139 29953: contig of 815 bp in length
* 29954 30053: gap of 100 bp
* 30054 30853: contig of 800 bp in length
* 30854 30953: gap of 100 bp
* 30954 31766: contig of 813 bp in length
* 31767 31866: gap of 100 bp
* 31867 32684: contig of 818 bp in length
* 32685 32784: gap of 100 bp
* 32785 33596: contig of 812 bp in length
* 33597 33696: gap of 100 bp
* 33697 34516: contig of 820 bp in length
* 34517 34616: gap of 100 bp
* 34617 35432: contig of 816 bp in length
* 35433 35532: gap of 100 bp
* 35533 36356: contig of 824 bp in length
* 36357 36456: gap of 100 bp
* 36457 37272: contig of 816 bp in length
* 37273 37372: gap of 100 bp
* 37373 38181: contig of 809 bp in length
* 38182 38281: gap of 100 bp
* 38282 39088: contig of 807 bp in length
* 39089 39999: contig of 811 bp in length
* 40000 40912: contig of 813 bp in length
* 40913 41012: gap of 100 bp
* 41013 41820: contig of 808 bp in length
* 41821 41920: gap of 100 bp
* 41921 42727: contig of 807 bp in length
* 42728 42827: gap of 100 bp
* 42828 43636: contig of 809 bp in length
* 43637 43736: gap of 100 bp
* 43737 44524: contig of 788 bp in length
* 44525 44625: gap of 100 bp
* 44626 45337: contig of 813 bp in length
* 45338 46343: contig of 806 bp in length
* 46344 46444: gap of 100 bp
* 46444 47250: contig of 807 bp in length
* 47251 47351: gap of 100 bp
* 47351 48152: contig of 802 bp in length
* 48153 48252: gap of 100 bp
* 48253 49066: contig of 814 bp in length
* 49067 49167: gap of 100 bp
* 49167 50086: contig of 820 bp in length
* 50087 50883: contig of 797 bp in length
* 50884 51769: gap of 100 bp
* 51770 51870: gap of 100 bp
* 51870 52666: contig of 797 bp in length
* 52667 53592: contig of 826 bp in length
* 53592 53692: gap of 100 bp
* 53693 54501: contig of 809 bp in length
* 54502 54602: contig of 809 bp in length
* 54603 55411: gap of 100 bp
* 55412 56327: contig of 817 bp in length
* 56328 56428: gap of 100 bp
* 56429 57237: contig of 810 bp in length
* 57238 58146: contig of 809 bp in length
* 58147 58247: gap of 100 bp
* 58248 59054: contig of 807 bp in length
* 59055 59154: gap of 100 bp
* 59155 59962: contig of 808 bp in length
* 59963 60062: gap of 100 bp
* 60063 60882: contig of 821 bp in length
* 60883 61792: gap of 100 bp
* 61793 61892: contig of 810 bp in length
* 61893 62703: contig of 811 bp in length
* 62704 62804: gap of 100 bp
* 62805 63616: contig of 812 bp in length
* 63617 63716: gap of 100 bp
* 63717 64537: contig of 821 bp in length
* 64538 65446: gap of 100 bp
* 65447 65546: contig of 810 bp in length
* 65547 65646: gap of 100 bp

Query Match 4.6%; Score 182; DB 14; Length 73656;
Best Local Similarity 100.0%; Pred. No. 6.6e-94;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1578 GGCACAGAAACAGACACAAAATCAGAGTGTCACCCGGGCTACATCTGCTTGAA 1637
Db 18798 GGCACAGAAACAGACACAAAATCAGAGTGTCACCCGGGCTACATCTGCTTGAA 18739

QY 1638 AAACTGGGCTTTAATCCAAATGACATATTTTGACCCCTAATCTTACCATGGGAC 1697
Db 18738 AAACTGGGCTTTAATCCAAATGACATATTTTGACCCCTAATCTTACCATGGGAC 18679

QY 1698 TGAATGGAGGACACAACTTGATGCCATTAATTTTATCCATGCAACAAAGTCATTA 1757
Db 18678 TGAATGGAGGACACAACTTGATGCCATTAATTTTATCCATGCAACAAAGTCATTA 18619

QY 1758 AG 1759
Db 18618 AG 18617

RESULT 31
BV204261
LOCUS sqm214563 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
DEFINITION tagged site.
ACCESSION BV204261
VERSION BV204261.1 GI:48173869
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 201)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
```

Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: NO primer sequence submitted  
Primer B: NO primer sequence submitted  
STS size: 201.

FEATURES  
source Location/Qualifiers  
1..201  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (sequenom)"  
<1..>201

STS  
ORIGIN

Query Match 3.3%; Score 129; DB 10; Length 201;  
Best Local Similarity 99.4%; Pred. No. 9.8e-63;  
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3560 TGGGTTACAGGCGATCCGCGGCTCTGGCTACCCAGCCAGCCGACACACCGAGA 3619  
|||||  
DB 22 TGGGTTACAGGCGATCCGCGGCTCTGGCTACCCAGCCAGCCGACACACCGAGA 81  
|||||

QY 3620 AGCTCACCATGTGAGACTCGACATCGAGAGTCTACAGGCAATTAGTTAACAGAT 3679  
|||||  
DB 82 AGCTCACCATGTGAGACTCGACATCGAGAGTCTACAGGCAATTAGTTAACAGAT 141  
|||||

QY 3680 CATTAGCAATGGCACCTGCTTACAGCAGTCTCAGCCCTCTACTTCTCCAAATTGAAGTCCA 3739  
|||||  
DB 142 CATTAGCAATGGCACCTGCTTACAGCAGTCTCAGCCCTCTACTTCTCCAAATTGAAGTCCA 201  
|||||

RESULT 32  
AC080113  
LOCUS  
DEFINITION AC080113 67611 bp DNA linear HTG 24-SEP-2000  
SAMPLING.  
AC080113  
VERSION AC080113.1 GI:10281431  
KEYWORDS HTG; HTGS PHASE0.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 67611)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 5, clone RP11-28J13  
Unpublished  
2 (bases 1 to 67611)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,  
Bouhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
Choeil,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sognuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Telamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (24-SEP-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RN/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L11103  
Center clone name: 28\_J\_13  
-----  
\* NOTE: This record contains 82 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 714: contig of 714 bp in length  
\* 715 814: gap of 100 bp  
\* 815 1528: contig of 714 bp in length  
\* 1529 1628: gap of 100 bp  
\* 1629 2367: contig of 739 bp in length  
\* 2368 2467: gap of 100 bp  
\* 2468 3202: contig of 735 bp in length  
\* 3203 3302: gap of 100 bp  
\* 3303 4039: contig of 737 bp in length  
\* 4040 4139: gap of 100 bp  
\* 4140 4875: contig of 736 bp in length  
\* 4876 4975: gap of 100 bp  
\* 4976 5685: contig of 710 bp in length  
\* 5686 6533: contig of 748 bp in length  
\* 6534 6634: gap of 100 bp  
\* 6634 7341: contig of 707 bp in length  
\* 7341 7441: contig of 717 bp in length  
\* 7441 8157: gap of 100 bp  
\* 8158 8257: contig of 716 bp in length  
\* 8258 8973: contig of 716 bp in length  
\* 8974 9073: gap of 100 bp  
\* 9074 9811: contig of 738 bp in length  
\* 9812 9912: gap of 100 bp  
\* 9912 10567: contig of 656 bp in length  
\* 10568 10667: gap of 100 bp  
\* 10668 11375: contig of 708 bp in length  
\* 11376 11475: gap of 100 bp  
\* 11476 12198: contig of 723 bp in length  
\* 12199 12298: gap of 100 bp  
\* 12299 13025: contig of 727 bp in length  
\* 13026 13125: gap of 100 bp  
\* 13126 13849: contig of 724 bp in length  
\* 13850 13949: gap of 100 bp  
\* 13950 14679: contig of 730 bp in length  
\* 14680 14779: gap of 100 bp  
\* 14780 15490: contig of 711 bp in length  
\* 15491 15590: gap of 100 bp  
\* 15591 16338: contig of 748 bp in length  
\* 16339 16439: gap of 100 bp  
\* 16439 17146: contig of 708 bp in length  
\* 17147 17246: gap of 100 bp  
\* 17247 17956: contig of 710 bp in length  
\* 17957 18056: gap of 100 bp  
\* 18057 18775: contig of 719 bp in length  
\* 18776 18876: gap of 100 bp  
\* 18876 19602: contig of 727 bp in length  
\* 19603 19702: gap of 100 bp  
\* 19703 20435: contig of 733 bp in length  
\* 20436 20535: gap of 100 bp

\* 20536 21263: contig of 728 bp in length  
\* 21264 21363: gap of 100 bp  
\* 21364 22084: contig of 721 bp in length  
\* 22085 22184: gap of 100 bp  
\* 22185 22915: contig of 731 bp in length  
\* 22916 23015: gap of 100 bp  
\* 23016 23745: contig of 730 bp in length  
\* 23746 23845: gap of 100 bp  
\* 23846 24577: contig of 732 bp in length  
\* 24578 25400: contig of 723 bp in length  
\* 25401 25500: gap of 100 bp  
\* 25501 26227: contig of 727 bp in length  
\* 26228 26327: gap of 100 bp  
\* 26328 27044: contig of 717 bp in length  
\* 27045 27145: gap of 100 bp  
\* 27145 27875: contig of 731 bp in length  
\* 27876 28690: gap of 100 bp  
\* 28691 28790: gap of 100 bp  
\* 28791 29521: contig of 731 bp in length  
\* 29522 29621: gap of 100 bp  
\* 29622 30352: contig of 731 bp in length  
\* 30353 30453: gap of 100 bp  
\* 30453 31189: contig of 737 bp in length  
\* 31190 31289: gap of 100 bp  
\* 31290 32015: contig of 726 bp in length  
\* 32016 32115: gap of 100 bp  
\* 32116 32847: contig of 732 bp in length  
\* 32848 32947: gap of 100 bp  
\* 32948 33671: contig of 724 bp in length  
\* 33672 33771: gap of 100 bp  
\* 33772 34501: contig of 730 bp in length  
\* 34502 35334: gap of 100 bp  
\* 35335 35434: gap of 100 bp  
\* 35435 36158: contig of 724 bp in length  
\* 36159 36258: gap of 100 bp  
\* 36259 37010: contig of 752 bp in length  
\* 37011 37110: gap of 100 bp  
\* 37111 37854: contig of 744 bp in length  
\* 37855 37954: gap of 100 bp  
\* 37955 38696: contig of 742 bp in length  
\* 38697 38796: gap of 100 bp  
\* 38797 39525: contig of 729 bp in length  
\* 39526 39625: gap of 100 bp  
\* 39626 40339: contig of 714 bp in length  
\* 40340 40439: gap of 100 bp  
\* 40440 41161: contig of 722 bp in length  
\* 41162 41261: gap of 100 bp  
\* 41262 41981: contig of 720 bp in length  
\* 41982 42081: gap of 100 bp  
\* 42082 42807: contig of 726 bp in length  
\* 42808 42907: gap of 100 bp  
\* 42908 43648: contig of 741 bp in length  
\* 43649 43748: gap of 100 bp  
\* 43749 44467: contig of 719 bp in length  
\* 44468 44567: gap of 100 bp  
\* 44568 45365: contig of 698 bp in length  
\* 45366 46081: gap of 100 bp  
\* 46082 46181: gap of 100 bp  
\* 46182 46908: contig of 727 bp in length  
\* 46909 47008: gap of 100 bp  
\* 47009 47739: contig of 731 bp in length  
\* 47740 47839: gap of 100 bp  
\* 47840 48565: contig of 726 bp in length  
\* 48566 49390: gap of 100 bp  
\* 48666 49391: contig of 725 bp in length  
\* 49391 49490: gap of 100 bp  
\* 49491 50216: contig of 726 bp in length  
\* 50217 50317: gap of 100 bp  
\* 50317 51006: contig of 690 bp in length

\* 51007 51106: gap of 100 bp  
\* 51107 51816: contig of 710 bp in length  
\* 51817 52656: gap of 100 bp  
\* 52657 52756: contig of 740 bp in length  
\* 52757 53479: gap of 100 bp  
\* 53480 53579: contig of 723 bp in length  
\* 53580 54315: gap of 100 bp  
\* 54316 54415: contig of 736 bp in length  
\* 54416 55154: gap of 100 bp  
\* 55155 55254: contig of 739 bp in length  
\* 55255 55985: gap of 100 bp  
\* 55986 56085: contig of 731 bp in length  
\* 56086 56811: gap of 100 bp  
\* 56812 56911: contig of 726 bp in length

Query Match 3.3%; Score 129; DB 14; Length 67611;  
Best Local Similarity 100.0%; Pred. No. 9.7e-63;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2913 AGTGAAGCCACGTTTATTGGGACCCAGCTTTTGAAGACTATGACCTGCAGAGCTGGT 2972  
DB 30644 AGTGAAGCCACGTTTATTGGGACCCAGCTTTTGAAGACTATGACCTGCAGAGCTGGT 30703

QY 2973 GGACTACATTGACTGGAAGCCCTTTTGTGCTGCGAGCTCCGGGGCAAGTACCGAA 3032  
DB 30704 GGACTACATTGACTGGAAGCCCTTTTGTGCTGCGAGCTCCGGGGCAAGTACCGAA 30763

QY 3033 TCGAGGCTT 3041  
DB 30764 TCGAGGCTT 30772

## RESULT 33

AC080113/c  
LOCUS AC080113  
DEFINITION Homo sapiens chromosome 5 clone RP11-28J13 map 5, linear HTG 24-SEP-2000  
SAMPLING.

AC080113  
VERSION AC080113.1 GI:10281431

KEYWORDS HTG; HTGS PHASE0.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

1 (bases 1 to 67611)

Birken,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 5, clone RP11-28J13

Unpublished

2 (bases 1 to 67611)

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Birken,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,E., Boguslavsky,L.,  
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,  
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
Lamazares,R., Landers,T., Lechoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McSwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neil,D., Olivat,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,K., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission

## TITLE

Submitted (24-SEP-2000) Whitehead Institute/MIT Center for Genome

## COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L11103

Center clone name: 28\_J\_13

-----  
 \* NOTE: This record contains 82 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 714: contig of 714 bp in length  
 \* 715 814: gap of 100 bp  
 \* 815 1528: contig of 714 bp in length  
 \* 1529 1628: gap of 100 bp  
 \* 1629 2367: contig of 739 bp in length  
 \* 2368 2467: gap of 100 bp  
 \* 2468 3202: contig of 735 bp in length  
 \* 3203 3302: gap of 100 bp  
 \* 3303 4039: contig of 737 bp in length  
 \* 4040 4139: gap of 100 bp  
 \* 4140 4875: contig of 736 bp in length  
 \* 4876 4975: gap of 100 bp  
 \* 4976 5685: contig of 710 bp in length  
 \* 5686 5785: gap of 100 bp  
 \* 5786 6533: contig of 748 bp in length  
 \* 6534 6633: gap of 100 bp  
 \* 6634 7340: contig of 707 bp in length  
 \* 7341 7440: gap of 100 bp  
 \* 7441 8157: contig of 717 bp in length  
 \* 8158 8257: gap of 100 bp  
 \* 8258 8973: contig of 716 bp in length  
 \* 8974 9073: gap of 100 bp  
 \* 9074 9811: contig of 738 bp in length  
 \* 9812 9911: gap of 100 bp  
 \* 9912 10567: contig of 656 bp in length  
 \* 10568 10667: gap of 100 bp  
 \* 10668 11375: contig of 708 bp in length  
 \* 11376 11475: gap of 100 bp  
 \* 11476 12198: contig of 723 bp in length  
 \* 12199 12298: gap of 100 bp  
 \* 12299 13025: contig of 727 bp in length  
 \* 13026 13125: gap of 100 bp  
 \* 13126 13849: contig of 724 bp in length  
 \* 13850 13949: gap of 100 bp  
 \* 13950 14679: contig of 730 bp in length  
 \* 14680 14779: gap of 100 bp  
 \* 14780 15490: contig of 711 bp in length  
 \* 15491 15590: gap of 100 bp  
 \* 15591 16338: contig of 748 bp in length  
 \* 16339 16438: gap of 100 bp  
 \* 16439 17146: contig of 708 bp in length  
 \* 17147 17246: gap of 100 bp  
 \* 17247 17956: contig of 710 bp in length  
 \* 17957 18056: gap of 100 bp  
 \* 18057 18775: contig of 719 bp in length  
 \* 18776 18875: gap of 100 bp  
 \* 18876 19602: contig of 727 bp in length  
 \* 19603 19702: gap of 100 bp  
 \* 19703 20435: contig of 733 bp in length

20436 20535: gap of 100 bp  
 \* 20536 21263: contig of 728 bp in length  
 \* 21264 21363: gap of 100 bp  
 \* 21364 22084: contig of 721 bp in length  
 \* 22085 22184: gap of 100 bp  
 \* 22185 22915: contig of 731 bp in length  
 \* 22916 23015: gap of 100 bp  
 \* 23016 23745: contig of 730 bp in length  
 \* 23746 23845: gap of 100 bp  
 \* 23846 24577: contig of 732 bp in length  
 \* 24578 24677: gap of 100 bp  
 \* 24678 25400: contig of 723 bp in length  
 \* 25401 25500: gap of 100 bp  
 \* 25501 26227: contig of 727 bp in length  
 \* 26228 26327: gap of 100 bp  
 \* 26328 27044: contig of 717 bp in length  
 \* 27045 27145: gap of 100 bp  
 \* 27145 27875: contig of 731 bp in length  
 \* 27876 27975: gap of 100 bp  
 \* 27976 28690: contig of 715 bp in length  
 \* 28691 28790: gap of 100 bp  
 \* 28791 29521: contig of 731 bp in length  
 \* 29522 29622: gap of 100 bp  
 \* 29622 30352: contig of 731 bp in length  
 \* 30353 30452: gap of 100 bp  
 \* 30453 31189: contig of 737 bp in length  
 \* 31190 31289: gap of 100 bp  
 \* 31290 32015: contig of 726 bp in length  
 \* 32016 32115: gap of 100 bp  
 \* 32116 32847: contig of 732 bp in length  
 \* 32848 32947: gap of 100 bp  
 \* 32948 33671: contig of 724 bp in length  
 \* 33672 33771: gap of 100 bp  
 \* 33772 34501: contig of 730 bp in length  
 \* 34502 34601: gap of 100 bp  
 \* 34602 35334: contig of 733 bp in length  
 \* 35335 35434: gap of 100 bp  
 \* 35435 36158: contig of 724 bp in length  
 \* 36159 36258: gap of 100 bp  
 \* 36259 37010: contig of 752 bp in length  
 \* 37011 37110: gap of 100 bp  
 \* 37111 37854: contig of 744 bp in length  
 \* 37855 37954: gap of 100 bp  
 \* 37955 38696: contig of 742 bp in length  
 \* 38697 38796: gap of 100 bp  
 \* 38797 39525: contig of 729 bp in length  
 \* 39526 39626: gap of 100 bp  
 \* 39626 40339: contig of 714 bp in length  
 \* 40340 40439: gap of 100 bp  
 \* 40440 41161: contig of 722 bp in length  
 \* 41162 41261: gap of 100 bp  
 \* 41262 41981: contig of 720 bp in length  
 \* 41982 42081: gap of 100 bp  
 \* 42082 42807: contig of 726 bp in length  
 \* 42808 42907: gap of 100 bp  
 \* 42908 43648: contig of 741 bp in length  
 \* 43649 43748: gap of 100 bp  
 \* 43749 44467: contig of 719 bp in length  
 \* 44468 44567: gap of 100 bp  
 \* 44568 45265: contig of 698 bp in length  
 \* 45266 45365: gap of 100 bp  
 \* 45366 46081: contig of 716 bp in length  
 \* 46082 46181: gap of 100 bp  
 \* 46182 46908: contig of 727 bp in length  
 \* 46909 47008: gap of 100 bp  
 \* 47009 47739: contig of 731 bp in length  
 \* 47740 47839: gap of 100 bp  
 \* 47840 48565: contig of 726 bp in length  
 \* 48566 48665: gap of 100 bp  
 \* 48666 49390: contig of 725 bp in length  
 \* 49391 49491: gap of 100 bp  
 \* 49491 50216: contig of 726 bp in length  
 \* 50217 50316: gap of 100 bp





```

* 57962 58061: gap of 100 bp
* 58062 59375: contig of 1314 bp in length
* 59376 59475: gap of 100 bp
* 59476 63078: contig of 3603 bp in length
* 63079 63178: gap of 100 bp
* 63179 65356: contig of 2178 bp in length
* 65357 65456: gap of 100 bp
* 65457 68593: contig of 3127 bp in length
* 68594 68683: gap of 100 bp
* 68684 73074: contig of 4391 bp in length
* 73075 73175: gap of 100 bp
* 73176 76775: contig of 3501 bp in length
* 76776 81036: contig of 4261 bp in length
* 81037 81136: gap of 100 bp
* 81137 86818: contig of 5682 bp in length
* 86819 86918: gap of 100 bp
* 86919 94761: contig of 7843 bp in length
* 94762 94862: gap of 100 bp
* 94863 102807: contig of 7945 bp in length
* 102808 102906: gap of 100 bp
* 102907 117954: contig of 15048 bp in length
* 117955 118054: gap of 100 bp
* 118055 130927: contig of 12873 bp in length
* 130928 131027: gap of 100 bp
* 131028 145686: contig of 14659 bp in length
* 145687 145687: gap of 100 bp
* 145688 162777: contig of 16991 bp in length
* 162778 162778: gap of 100 bp
* 162779 164529: contig of 1652 bp in length.
FEATURES             Location/Qualifiers
     source            1..164529
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="1"
                        /map="1"
                        /clone="RP11-11N21"
misc_feature          1..638
                        /note="assembly_fragment"
                        /clone_end="SP6"
                        /vector_side="left"
gap                  639..738
                        /estimated_length=100
misc_feature          739..1280
                        /note="assembly_fragment"
gap                  1281..1380
                        /estimated_length=100
misc_feature          1381..2021
                        /note="assembly_fragment"
gap                  2022..2121
                        /estimated_length=100
misc_feature          2122..2830
                        /note="assembly_fragment"
gap                  2831..2930
                        /estimated_length=100
misc_feature          2931..3970
                        /note="assembly_fragment"
gap                  3971..4070
                        /estimated_length=100
misc_feature          4071..4923
                        /note="assembly_fragment"
gap                  4924..5023
                        /estimated_length=100
misc_feature          5024..6063
                        /note="assembly_fragment"
gap                  6064..6163
                        /estimated_length=100
misc_feature          6164..7446
                        /note="assembly_fragment"
gap                  7447..7546
                        /estimated_length=100

```

```

misc_feature          7547..9381
                        /note="assembly_fragment"
gap                  9382..9481
                        /estimated_length=100
misc_feature          9482..11219
                        /note="assembly_fragment"
gap                  11220..11319
                        /estimated_length=100
misc_feature          11320..12350
                        /note="assembly_fragment"
gap                  12351..12450
                        /estimated_length=100
misc_feature          12451..13291
                        /note="assembly_fragment"

Query Match          2.9%; Score 114; DB 14; Length 164529;
Best Local Similarity 100.0%; Pred. No. 6.4e-54;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2255 CTCAGGTTATAAGTCAGCCGGGTTATGAAGAAGGCTTGGCCACCTTATCCCTTTCA 2314
Db 87313 CTCAGGTTATAAGTCAGCCGGGTTATGAAGAAGGCTTGGCCACCTTATCCCTTTCA 87372

QY 2315 TCGAAAAGAAAGAGAGAAACACAGAGTCTTAAACGGCACAGTAGAAGAGAGG 2368
Db 87373 TCGAAAAGAAAGAGAGAAACACAGAGTCTTAAACGGCACAGTAGAAGAGAGG 87426

RESULT 35
AC016512              73656 bp DNA linear HTG 13-JUL-2000
LOCUS                Homo sapiens clone RP11-11N21, LOW-PASS SEQUENCE SAMPLING.
DEFINITION            AC016512
ACCESSION              AC016512.3 GI:9123838
VERSION                HTG; HTGS_PHASE0.
KEYWORDS                Homo sapiens (human)
SOURCE                 Homo sapiens
ORGANISM                Homo sapiens
                        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                        Hominidae; Homo.
REFERENCE              1 (bases 1 to 73656)
AUTHORS                Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE                  Homo sapiens, clone RP11-11N21
JOURNAL                Unpublished
REFERENCE              2 (bases 1 to 73656)
AUTHORS                Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
                        Baldwin,J., Barna,N., Beckerly,K., Boguslavsky,L., Bouckhalter,B.,
                        Brown,A., Castle,A., Collings,M., Collins,S., Collymore,A.,
                        Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
                        Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
                        Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
                        Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
                        Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,
                        McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
                        Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
                        Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
                        Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
                        Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
                        Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
                        Direct Submission
                        Submitted (01-DEC-1999) Whitehead Institute/MIT Center for Genome
                        Research, 320 Charles Street, Cambridge, MA 02141, USA
                        On Jul 13, 2000 this sequence version replaced gi:7144825.
                        All repeats were identified using RepeatMasker:
                        Smit, A.F.A. & Green, P. (1996-1997)
                        http://ftp.genome.washington.edu/RM/RepeatMasker.html
                        ----- Genome Center
                        Center: Whitehead Institute/ MIT Center for Genome Research
                        Center code: WIBR
                        Web site: http://www-seq.wi.mit.edu
                        Contact: sequence_submissions@genome.wi.mit.edu
                        ----- Project Information
                        Center project name: L3233
                        Center clone name: 11_N_21

```

-----  
\* NOTE: This record contains 81 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 809: contig of 809 bp in length  
\* 810 909: gap of 100 bp  
\* 910 1718: contig of 809 bp in length  
\* 1719 1818: gap of 100 bp  
\* 1819 2635: contig of 817 bp in length  
\* 2636 2735: gap of 100 bp  
\* 2736 3541: contig of 806 bp in length  
\* 3542 3641: gap of 100 bp  
\* 3642 4430: contig of 789 bp in length  
\* 4431 4530: gap of 100 bp  
\* 4531 5317: contig of 787 bp in length  
\* 5318 5417: gap of 100 bp  
\* 5418 6230: contig of 813 bp in length  
\* 6231 6330: gap of 100 bp  
\* 6331 7135: contig of 805 bp in length  
\* 7136 7235: gap of 100 bp  
\* 7236 8055: contig of 820 bp in length  
\* 8056 8155: gap of 100 bp  
\* 8156 8959: contig of 804 bp in length  
\* 8960 9059: gap of 100 bp  
\* 9060 9873: contig of 814 bp in length  
\* 9874 9974: gap of 100 bp  
\* 9974 10777: contig of 804 bp in length  
\* 10778 10877: gap of 100 bp  
\* 10878 11696: contig of 819 bp in length  
\* 11697 11797: gap of 100 bp  
\* 11797 12593: contig of 797 bp in length  
\* 12594 12693: gap of 100 bp  
\* 12694 13490: contig of 797 bp in length  
\* 13491 13590: gap of 100 bp  
\* 13591 14402: contig of 812 bp in length  
\* 14403 14502: gap of 100 bp  
\* 14503 15304: contig of 802 bp in length  
\* 15305 15405: gap of 100 bp  
\* 15405 16210: contig of 806 bp in length  
\* 16211 16310: gap of 100 bp  
\* 16311 17114: contig of 804 bp in length  
\* 17115 17214: gap of 100 bp  
\* 17215 18042: contig of 828 bp in length  
\* 18043 18142: gap of 100 bp  
\* 18143 18958: contig of 816 bp in length  
\* 18959 19058: gap of 100 bp  
\* 19059 19870: contig of 812 bp in length  
\* 19871 19970: gap of 100 bp  
\* 19971 20775: contig of 805 bp in length  
\* 20776 20875: gap of 100 bp  
\* 20876 21691: contig of 816 bp in length  
\* 21692 21791: gap of 100 bp  
\* 21792 22615: contig of 824 bp in length  
\* 22616 22715: gap of 100 bp  
\* 22716 23552: contig of 837 bp in length  
\* 23553 23652: gap of 100 bp  
\* 23653 24462: contig of 810 bp in length  
\* 24463 24562: gap of 100 bp  
\* 24563 25381: contig of 819 bp in length  
\* 25382 25481: gap of 100 bp  
\* 25482 26299: contig of 818 bp in length  
\* 26300 26399: gap of 100 bp  
\* 26400 27217: contig of 818 bp in length  
\* 27317 27318: gap of 100 bp  
\* 27318 28122: contig of 805 bp in length  
\* 28123 28222: gap of 100 bp  
\* 28223 29038: contig of 816 bp in length  
\* 29039 29138: gap of 100 bp  
\* 29139 29953: contig of 815 bp in length  
\* 29954 30053: gap of 100 bp  
\* 30054 30853: contig of 800 bp in length  
\* 30854 30953: gap of 100 bp  
\* 30954 31766: contig of 813 bp in length  
\* 31767 31866: gap of 100 bp  
\* 31867 32685: contig of 818 bp in length  
\* 32685 32785: gap of 100 bp  
\* 32785 33597: contig of 812 bp in length  
\* 33597 34516: gap of 100 bp  
\* 34517 34617: contig of 820 bp in length  
\* 34617 35432: gap of 100 bp  
\* 35432 35532: contig of 816 bp in length  
\* 35533 36356: gap of 100 bp  
\* 36356 36456: contig of 824 bp in length  
\* 36457 37272: gap of 100 bp  
\* 37273 37372: contig of 816 bp in length  
\* 37372 38181: gap of 100 bp  
\* 38182 38281: contig of 809 bp in length  
\* 38282 39088: contig of 807 bp in length  
\* 39089 39188: gap of 100 bp  
\* 39189 39999: contig of 811 bp in length  
\* 40000 40912: gap of 100 bp  
\* 40913 41012: contig of 813 bp in length  
\* 41013 41820: contig of 808 bp in length  
\* 41821 41920: gap of 100 bp  
\* 41921 42727: contig of 807 bp in length  
\* 42728 43636: gap of 100 bp  
\* 43637 43736: contig of 809 bp in length  
\* 43737 44244: gap of 100 bp  
\* 44245 44625: contig of 788 bp in length  
\* 44626 45337: gap of 100 bp  
\* 45338 45438: contig of 813 bp in length  
\* 45439 46343: gap of 100 bp  
\* 46344 46444: contig of 806 bp in length  
\* 46445 47250: gap of 100 bp  
\* 47251 47350: contig of 807 bp in length  
\* 47351 48152: gap of 100 bp  
\* 48153 48252: contig of 802 bp in length  
\* 48253 49066: gap of 100 bp  
\* 49067 49166: contig of 814 bp in length  
\* 49167 49886: gap of 100 bp  
\* 49887 50086: contig of 820 bp in length  
\* 50087 50883: gap of 100 bp  
\* 50884 51769: contig of 797 bp in length  
\* 50984 51769: gap of 100 bp  
\* 51770 51869: contig of 786 bp in length  
\* 51870 52666: gap of 100 bp  
\* 52667 52766: contig of 797 bp in length  
\* 52767 53592: gap of 100 bp  
\* 53593 53693: contig of 826 bp in length  
\* 53693 54501: gap of 100 bp  
\* 54502 54601: contig of 809 bp in length  
\* 54602 55410: gap of 100 bp  
\* 55411 55510: contig of 809 bp in length  
\* 55511 56327: gap of 100 bp  
\* 56328 56428: contig of 817 bp in length  
\* 56429 57237: gap of 100 bp  
\* 57238 57337: contig of 810 bp in length  
\* 57338 58146: gap of 100 bp  
\* 58147 58247: contig of 809 bp in length  
\* 58248 59054: gap of 100 bp  
\* 59054 59154: contig of 807 bp in length  
\* 59154 59961: gap of 100 bp  
\* 59962 60061: contig of 808 bp in length  
\* 60062 60882: gap of 100 bp  
\* 60883 60982: contig of 821 bp in length  
\* 60983 60983: gap of 100 bp

\* 60983 61792: contig of 810 bp in length  
 \* 61793 61892: gap of 100 bp  
 \* 61893 62703: contig of 811 bp in length  
 \* 62704 62803: gap of 100 bp  
 \* 62804 63615: contig of 812 bp in length  
 \* 63616 63715: gap of 100 bp  
 \* 63716 64536: contig of 821 bp in length  
 \* 64537 64636: gap of 100 bp  
 \* 64637 65446: contig of 810 bp in length  
 \* 65447 65546: gap of 100 bp

Query Match 2.4%; Score 103; DB 14; Length 73656;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-47;  
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2366 AGACCCCTTACGAGGACCATCGTGTGCGCCACTGTTAAAGCGGACGTCGACGACATAG 2425  
 DB 32416 AGGACCCCTTACGAGGACCATCGTGTGCGCCACTGTTAAAGCGGACGTCGACGACATAG 32475

QY 2426 GCAAGACATAGTTGGAGTAGTCTTGGCTGCAATAATTTCCG 2468  
 DB 32476 GCAAGACATAGTTGGAGTAGTCTTGGCTGCAATAATTTCCG 32518

RESULT 36  
 LOCUS CQ750107/c 732 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 36041 from Patent WO02068579.

ACCESSION CQ750107  
 VERSION CQ750107.1 GI:42380503  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kites, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 02068579-A 36041 06-SEP-2002;  
 PE Corporation (NY) (US)

FEATURES  
 source  
 1. .732  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 2.4%; Score 96; DB 6; Length 732;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-43;  
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTACCTGTGAGAGCAGCTTCTCTGCGCGCCCTCTGCGAAGGAGGAGCTCGACA 61  
 DB 267 GTACCTGTGAGAGCAGCTTCTCTGCGCGCCCTCTGCGAAGGAGGAGCTCGACA 208

QY 62 ACATGTACCCCGCGCTCCAGACCTGTGCGAACCCG 97  
 DB 207 ACATGTACCCCGCGCTCCAGACCTGTGCGAACCCG 172

RESULT 37  
 LOCUS BV599990 731 bp DNA linear STS 14-APR-2005  
 DEFINITION S217P60970RH2.T0 Noemie Pan troglodytes troglodytes STS genomic, sequence tagged site.

ACCESSION BV599990  
 VERSION BV599990.1 GI:62538904  
 KEYWORDS STS.  
 SOURCE Pan troglodytes troglodytes  
 ORGANISM Pan troglodytes troglodytes

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Pan. 1 to 731)  
 Mikkelsen, T.S., Hillier, W.L., Eichler, E.B., Zody, M.C. and  
 Jaffe, D.B.  
 Initial Sequence of the Chimpanzee Genome and Comparison with the  
 Human Genome  
 Unpublished (2005)

Contact: Michael C. Zody  
 Broad Institute of MIT and Harvard  
 320 Charles Street, Cambridge, MA 02141, USA  
 Tel: 6172580933  
 Fax: 6172580903  
 Email: mczody@broad.mit.edu  
 Primer A: No sequence submitted  
 Primer B: No sequence submitted  
 STS size: 731

Protocol:  
 23,021,928 chimpanzee whole genome shotgun reads were aligned to  
 the Human genome NCBI  
 Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,  
 including Clint (Pan  
 troglodytes verus), 3 other Pan troglodytes verus chimps  
 (Donald, Karlien, Yvonne), 3 Pan  
 troglodytes troglodytes chimps (Noemie, Masuku, Clara) and 2 chimps  
 of unknown origin  
 (Gon, Unknown Chimp). Common names: Pan troglodytes verus is the  
 western chimp and Pan  
 troglodytes troglodytes is the central chimp. To be included in  
 chimpanzee SNP discovery, a  
 read must be at least 500bp in length, at least 50% of its base  
 calls must have Phred  
 score >= 20, at least 30% of its base calls must satisfy  
 SNQS(30,25) (single strand NQS, the  
 base in question has Phred score >= 30, the surrounding 10 bases in  
 the read have Phred  
 score >= 25), and the read must have at least 200 bp SNQS(30,25)  
 bases. Reads not uniquely  
 placed in the genome and read pairs whose two ends were not  
 consistently placed were  
 discarded. After above filtering, NQS(30,25) standard was applied  
 to all pairs of  
 overlapping reads to call NQS bases and SNPs. Alignments (between  
 two reads) with less  
 than 100 NQS bases or with SNP rate > 0.01 were discarded. To  
 exclude alignment between two  
 copies of a single read, comparisons between two reads that share  
 95% of their genome  
 alignments (>=95% bases of read A and >=95% bases of read B were  
 placed at the same locus  
 of human genome) were discarded.

## FEATURES

source  
 1. .731  
 /organism="Pan troglodytes troglodytes"  
 /mol\_type="genomic DNA"  
 /sub\_species="troglodytes"  
 /db\_xref="taxon:37011"  
 /clone\_lib="Noemie"  
 <1. .>731

## ORIGIN

Query Match 2.4%; Score 95; DB 10; Length 731;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-43;  
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 733 ATTTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGACAGACGAGAGGGA 792  
 DB 412 ATTTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGACAGACGAGAGGGA 471  
 QY 793 TTTGTCTATCAGCGTGTCTCATGGAGAACCACTCTG 827

Db 472 TTGTCATCAGCTGTCTCATGGAGACCACTCTG 506

RESULT 38  
BV597233/c  
LOCUS  
DEFINITION S216P61129FGA.TO Masuku Pan troglodytes troglodytes STS 14-APR-2005  
sequence tagged site.

ACCESSION BV597233  
VERSION BV597233.1 GI:62536147  
KEYWORDS STS.  
SOURCE Pan troglodytes troglodytes  
ORGANISM Pan troglodytes troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Pan.  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and  
Jaffe,D.B.  
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the  
Human Genome  
JOURNAL Unpublished (2005)  
COMMENT  
Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 758  
Protocol:  
23,021,928 chimpanzee whole genome shotgun reads were aligned to  
the Human genome NCBI  
Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,  
including Clint (Pan  
troglodytes verus), 3 other Pan troglodytes verus chimps  
(Donald,Karlén,Yvonne), 3 Pan  
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps  
of unknown origin  
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the  
western chimp and Pan  
troglodytes troglodytes is the central chimp. To be included in  
chimpanzee SNP discovery, a  
read must be at least 500bp in length, at least 50% of its base  
calls must have Phred  
score >= 20, at least 30% of its base calls must satisfy  
SNQS(30,25) (single strand NQS, the  
base in question has Phred score >= 30, the surrounding 10 bases in  
the read have Phred  
score >= 25), and the read must have at least 200 bp SNQS(30,25)  
bases. Reads not uniquely  
placed in the genome and read pairs whose two ends were not  
consistently placed were  
discarded. After above filtering, NQS(30,25) standard was applied  
to all pairs of  
overlapping reads to call NQS bases and SNPs. Alignments (between  
two reads) with less  
than 100 NQS bases or with SNP rate > 0.01 were discarded. To  
exclude alignment between two  
copies of a single read, comparisons between two reads that share  
95% of their genome  
alignments (>=95% bases of read A and >=95% bases of read B were  
placed at the same locus  
of human genome) were discarded.  
Location/Qualifiers  
1. .758  
/organism="Pan troglodytes troglodytes"  
/mol\_type="genomic DNA"  
/sub\_species="troglodytes"  
/db\_xref="taxon:37011"  
/clone\_lib="Masuku"

ORIGIN  
Query Match 2.4%; Score 93; DB 10; Length 758;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 311 AGGAATACCTGCTGGCTGGGGCAGATATCATTTGAAACAATACTTTTAGCAGCACTAGTA 370  
DB 427 AGGAATACCTGCTGGCTGGGGCAGATATCATTTGAAACAATACTTTTAGCAGCACTAGTA 368  
QY 371 TTGCCCAAGCTGACTATGCTTGAACACTTGG 403  
DB 367 TTGCCCAAGCTGACTATGCTTGAACACTTGG 335

RESULT 39  
LOCUS  
DQ084519  
DEFINITION Bos taurus 5-methyltetrahydrofolate-homocysteine methyltransferase  
(MTR) mRNA, complete cds.  
ACCESSION DQ084519  
VERSION DQ084519.1 GI:68160374  
KEYWORDS Bos taurus (cow)  
SOURCE Bos taurus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 3832)  
AUTHORS Palin,M.F., Beaudry,D., Charest,R. and Girard,C.  
TITLE Interactions of folic acid-vitamin B12-methionine: effects on liver  
metabolism and production of dairy cows  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3832)  
AUTHORS Palin,M.F., Beaudry,D., Charest,R. and Girard,C.  
TITLE Direct Submission  
JOURNAL Submitted (03-JUN-2005) Dairy and Swine Research and Development  
Centre, P.O. Box 90, 2000 Route 108 East, Lennoxville, Quebec J1M  
1Z3, Canada

FEATURES  
source  
1. .3832  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="liver"  
1. .3832  
/genes="MTR"  
22. .3819  
/notes="methionine synthase"  
/codon\_start=1  
/product="5-methyltetrahydrofolate-homocysteine  
methyltransferase"  
/protein\_id="AA86762.1"  
/db\_xref="GI:68160375"

IDWPKPFDVQVLRGKYNRGPPIKIFDDKTGVBKAKVYDDAQNMLQALISQKQLQARG  
VGFMPAQSIODDHLYAEGAVQASPIATFYGLRQAAEKDSASDPYLCISGDFIAP  
LHSGIPDYGLFPAVACFGEELSAYEECCDDYSSIMVKALGDLAEAPAEHELHAR  
RELWGYCSGQLAVADURRKYRIGIRPAPGTPSPDHTEKLTWRILADVEORTGIRLT  
ESLAMAPASAVSGLYFSLNLSKYFAVGKISKQOIEDYASRKNMSVAEVEKWLPIILGY  
DTD"

variation

248 /gene="MTR"  
/replace="g"

variation

409 /gene="MTR"  
/replace="a"

variation

1552 /gene="MTR"  
/replace="a"

variation

3455 /gene="MTR"  
/replace="c"

ORIGIN

Query Match 1.7%; Score 68; DB 4; Length 3832;  
Best Local Similarity 100.0%; Pred. No. 7.3e-27;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2524 GACCACAAAGCAGATATAATTGGCTGTTCAGGACTCATCTCTCCCTGGATGAATG 2583  
|||||  
Db 2482 GACCACAAAGCAGATATAATTGGCTGTTCAGGACTCATCTCTCCCTGGATGAATG 2541  
|||||

Qy 2584 ATTTTGT 2591  
|||||Db 2542 ATTTTGT 2549  
|||||

RESULT 40

AF034214 4498 bp mRNA linear ROD 20-JAN-1999  
LOCUS Rattus norvegicus methionine synthase (MS) mRNA, complete cds.  
DEFINITION AF034214  
ACCESSION AF034214  
VERSION AF034214.1 GI:3978142

KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 4498)  
Yamada, K., Tobimatsu, T. and Toraya, T.  
Cloning, sequencing, and heterologous expression of rat methionine  
synthase cDNA

JOURNAL

BioSci. Biotechnol. Biochem. 62 (11), 2155-2160 (1998)

PUBMED

9972236

REFERENCE

2 (bases 1 to 4498)  
Yamada, K., Tobimatsu, T. and Toraya, T.  
Direct Submission  
Submitted (11-NOV-1997) Bioscience and Biotechnology, Okayama  
University, 3-1-1 Taishima-Naka, Okayama, Okayama 700, Japan

FEATURES

source

1. .4498  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/tissue\_type="liver"  
284. .4045  
/EC\_number="2.1.1.13"  
/note="MS"  
/codon\_start=1  
/product="methionine synthase"  
/protein\_id="AAD05384.1"  
/db\_xref="GI:3978143"

CDS

translation="MKTKLQDEIBAILKRIMVLDGGMGTMIRYKLSSENFQGEPEK  
DHSRPLKGNNDILSIYQIHKEYLLAGADIETITFSSTIAQADYGLSHLAY  
RMNKSADVARKAAEETLQGVKRFVAGSLGPTNKTLSVSPSPVERPDYRNITFDLV

EAYOEQAKGLDGVDDILLIETIFDTANAKAALFALQKLFEEYASPRPIFISGTIVD  
KSGRTSGOTGEAFVTSVSHSDPLCIGLNCALCAEMRPFETIGCTAYVLCYRNA  
GLPNTFGDDETAMAMHLDKDFAVDGLVNVVGGCGSTPDHIREAFAVCKPRVP  
PSVFECHMLLSGLEPFRPGPTNFVNIIGERCNVAGSKKFAKLIMAGNYEEALSVAK  
QVEMGAQVDDINDDGLDGPSTAKFCNFIASEPDIKVPCLCIDSNTFAVIEAGLKC  
COGKCIIVSISLKEGEEDELEKARKIKFGAAVNVAFDEEGQATETDKVSVCTRAY  
HLLVEKVGFPNDIIPDPNLTIGTGBEENLYAINFIHATRVIKETLFGVRISSGLS  
NLSFAPRGMDAIREAMHGVFLVHAIFKFGMDGIVNAGSLPVYDDIHKLQLCEDLIW  
NDAEATEKLYAVAOHKGKVKVLOTDEWRNGSIEERLEYALVKGIHVEDTBEA  
KERBEARVLNGSVEEEDPYGTTIVLATVGDVHDIGKNI VGVVLCGNFRVVIDGVMT  
PCDKILOALADHKADIGLSGLITPSLDEMI FVAKEMERLAIKIPILIGATTSTRT  
AVKIAPRYGAPVTHVLDASKSVVCSQLDENLKDQDYFEIILEEYEDIRQDHYESLKE  
RKYLPISQARKSHFDLWSEPHVKTPTFIGTQVPEDYNLQKLVYIDWKVPFVWOL  
RGKYNRGPKIFNDKAVGEARKVVEDAQNMLSLIISKKLRARGVGVGMPAQSVD  
DIHLYAEGAVQAAEFIAFYGLRQAAEKDSSTDPYHCLSFVALEHSGVRDYLGLF  
AVACFGVEELSAYEEDGDDYSIMVKALGDLAEAPAEHELHVRERRELLWAYCGSQL  
GVTDLRLRYEGIRPAPGTPSPDHTEKLTWRILANIEQATIRGIRLETSLAMAPASAVS  
GLYFSNVKSKYFAVGKISKQOIEDYALRKNMSVAEVEKWLGPILGYDTD"

ORIGIN

Query Match 1.7%; Score 65; DB 9; Length 4498;  
Best Local Similarity 100.0%; Pred. No. 4.2e-25;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2890 GATTGGCTGTCTCAACTCACCAGTGAAGCCACCGTTTATTGGGACCCAGGTCTTTGAA 2949  
|||||  
Db 3074 GATTGGCTGTCTCAACTCACCAGTGAAGCCACCGTTTATTGGGACCCAGGTCTTTGAA 3133  
|||||

Qy 2950 GACTA 2954  
|||||Db 3134 GACTA 3138  
|||||

RESULT 41

AX611831 63 bp DNA linear PAT 17-FEB-2003  
LOCUS Sequence 2856 from Patent WO02072882.  
DEFINITION AX611831  
ACCESSION AX611831  
VERSION AX611831.1 GI:28407260

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE

1  
Cullen, P. and Seedorf, U.  
Coronary chip  
Patent: WO 02072882-A 2856 19-SEP-2002;  
OSHAM GmbH (DE)

FEATURES

source

1. .63  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 1.6%; Score 63; DB 6; Length 63;  
Best Local Similarity 100.0%; Pred. No. 6.4e-24;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3544 GCAGACCTCGAAGGTTGCGGTACAAGGCATCCGCCGCTCTCGCTACCCCGACG 3603  
|||||  
Db 1 GCAGACCTCGAAGGTTGCGGTACAAGGCATCCGCCGCTCTCGCTACCCCGACG 60  
|||||

Qy 3604 CCC 3606  
|||||Db 61 CCC 63  
|||||

RESULT 42

AC156661/c



LOCUS	AC156661	211299 bp	DNA	linear	HTG 01-JUL-2005
DEFINITION	Bos taurus clone CH240-66D4, WORKING DRAFT SEQUENCE, 19 unordered pieces.				
ACCESSION	AC156661				
VERSION	AC156661.2 GI:68265967				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Bos taurus (cow)				
ORGANISM	Bos taurus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.				
AUTHORS	Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howles, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Loulseghe, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Popper, P., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojag, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.				
TITLE	Direct Submission				
REFERENCE	Unpublished				
JOURNAL	2 (bases 1 to 211299)				
AUTHORS	Worley, K. C.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 211299)				
AUTHORS	Cow Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-JUL-2005) Human Genome Sequencing Center, Department				

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Center: Baylor College of Medicine
On Jun 28, 2005 this sequence version replaced gi:58531381.	Center code: BCM
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: FDWA Center clone name: CH240-66D4 ----- Summary Statistics Assembly program: Atlas 3.0; Consensus quality: 203895 bases at least Q40 Consensus quality: 205492 bases at least Q30 Consensus quality: 206916 bases at least Q20 Estimated insert size: 206384; sum-of-contigs estimation Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
***** NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html). ***** NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. ***** This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	1770 1769: contig of 1769 bp in length 2194: gap of 425 bp 31094: contig of 28900 bp in length 31095 31144: gap of 50 bp 31145 45383: contig of 14239 bp in length 45384 45433: gap of 50 bp 45434 55239: contig of 9866 bp in length 55300 55349: gap of 50 bp 55350 84038: contig of 28689 bp in length 84039 109833: contig of 25745 bp in length 109834 109883: gap of 50 bp 109884 138955: contig of 29072 bp in length 138956 139005: gap of 50 bp 139006 149048: contig of 10043 bp in length 149049 149098: gap of 50 bp 149099 161796: contig of 12698 bp in length 161797 161846: gap of 50 bp 161847 194034: contig of 32178 bp in length 194035 194075: gap of 50 bp 194076 195658: contig of 1584 bp in length 195659 195758: gap of unknown length 195759 199096: contig of 3338 bp in length 199097 200578: gap of 1482 bp 200579 202040: contig of 1462 bp in length 202041 202140: gap of unknown length 202141 203220: contig of 1080 bp in length 203221 203320: gap of unknown length 203321 204553: contig of 1233 bp in length 204554 204653: gap of unknown length 204654 206247: contig of 1594 bp in length 206248 206347: gap of unknown length

```

* 206348 207768: contig of 1421 bp in length
* 207769 207868: gap of unknown length
* 207869 209529: contig of 1661 bp in length
* 209530 209629: gap of unknown length
* 209630 211299: contig of 1670 bp in length.
      Location/Qualifiers
        1. .211299
          /organism="Bos taurus"
          /mol_type="genomic DNA"
          /db_xref="taxon:9913"
          /clone="CH240-66D4"
          1770. .2194
            /estimated_length=425
          31095. .31144
            /estimated_length=50
          45384. .45433
            /estimated_length=50
          55300. .55349
            /estimated_length=50
          84039. .84088
            /estimated_length=50
          109834. .109883
            /estimated_length=50
          138956. .139005
            /estimated_length=50
          149049. .149098
            /estimated_length=50
          161797. .161846
            /estimated_length=50
          194025. .194074
            /estimated_length=50
          195659. .195758
            /estimated_length=unknown
          199097. .200578
            /estimated_length=1482
          202041. .202140
            /estimated_length=unknown
          203221. .203320
            /estimated_length=unknown
          204554. .204653
            /estimated_length=unknown
          206248. .206347
            /estimated_length=unknown
          207769. .207868
            /estimated_length=unknown
          209530. .209629
            /estimated_length=unknown
          /estimated_length=unknown

      ORIGIN
      Query Match      1.5%; Score 57; DB 14: Length 211299;
      Best Local Similarity 100.0%; Pred. No. 2.1e-20;
      Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      Qy 2535 AGATATAATGGCTGTGTCAGGACCTCATCTCTCCCTGGATGAATGATTTTGT 2591
          |||||
      Db 204259 AGATATAATGGCTGTGTCAGGACCTCATCTCTCCCTGGATGAATGATTTTGT 204203

      RESULT 43
      AX611827 55 bp DNA linear PAT 17-FEB-2003
      LOCUS
      DEFINITION Sequence 2852 from Patent WO02072882.
      ACCESSION AX611827
      VERSION AX611827.1 GI:28407256
      KEYWORDS
      SOURCE
      ORGANISM Homo sapiens (human)
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Hominidae; Homo.
      1
      Cullen, P. and Seedorf, U.
      Coronary chip

      REFERENCE
      AUTHORS
      TITLE
      JOURNAL

      Query Match      1.4%; Score 55; DB 6; Length 55;
      Best Local Similarity 100.0%; Pred. No. 3.2e-19;
      Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      Qy 2785 GAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAAGG 2839
          |||||
      Db 1 GAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAAGG 55

      RESULT 44
      AX611829 55 bp DNA linear PAT 17-FEB-2003
      LOCUS
      DEFINITION Sequence 2854 from Patent WO02072882.
      ACCESSION AX611829
      VERSION AX611829.1 GI:28407258
      KEYWORDS
      SOURCE
      ORGANISM Homo sapiens (human)
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
      Hominidae; Homo.
      1
      Cullen, P. and Seedorf, U.
      Coronary chip

      REFERENCE
      AUTHORS
      TITLE
      JOURNAL

      Query Match      1.4%; Score 55; DB 6; Length 55;
      Best Local Similarity 100.0%; Pred. No. 3.2e-19;
      Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      Qy 2785 GAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAAGG 2839
          |||||
      Db 1 GAAATCATGGAAGATATGAAGATATTAGACAGGACCATTTAGTCTCTCAAGG 55

      RESULT 45
      AC154523/c 194232 bp DNA linear HTG 12-JAN-2005
      LOCUS
      DEFINITION Mus musculus chromosome 13 clone R24-65H12, WORKING DRAFT
      SEQUENCE 9 unordered pieces.
      ACCESSION AC154523 AC133884
      VERSION AC154523.1 GI:56900149
      KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.
      SOURCE Mus musculus (house mouse)
      ORGANISM Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
      Sciurognathi; Muridae; Murinae; Mus.
      1 (bases 1 to 194232)
      Wilson, R.K.
      The sequence of Mus musculus clone
      Unpublished
      2 (bases 1 to 194232)
      Wilson, R.K.
      Direct Submission
      Submitted (30-DEC-2004) Genome Sequencing Center, 4444 Forest Park
      Parkway, St. Louis, MO 63108, USA
      JOURNAL
  
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 194232)  
Wilson, R.K.  
Direct Submission  
Submitted (12-JAN-2005) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On Dec 30, 2004 this sequence version replaced gi:28894649.

COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
----- Project Information -----  
Center project name: M\_BB0065H12  
Drafting center: WIBR

----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-terminator; 100%  
Chemistry: Dye-terminator; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 190157 bases at least Q40  
Consensus quality: 191134 bases at least Q30  
Consensus quality: 191685 bases at least Q20

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1179: contig of 1179 bp in length  
\* 1180: gap of unknown length  
\* 1280 2449: contig of 1170 bp in length  
\* 2450 2549: gap of unknown length  
\* 2550 4084: contig of 1535 bp in length  
\* 4085 4184: gap of unknown length  
\* 4185 6466: contig of 2282 bp in length  
\* 6467 6566: gap of unknown length  
\* 6567 8700: contig of 2134 bp in length  
\* 8701 8800: gap of unknown length  
\* 8801 24080: contig of 15280 bp in length  
\* 24081 24180: gap of unknown length  
\* 24181 54114: contig of 29934 bp in length  
\* 54115 54214: gap of unknown length  
\* 54215 87114: contig of 32900 bp in length  
\* 87115 87215: gap of unknown length  
\* 87215 194232: contig of 107018 bp in length.

FEATURES  
source

1. 194232  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="13"  
/clone="RP24-65H12"

misc\_feature 1. 1179  
/note="assembly\_name:Contig24"  
gap 1180. 1279  
/estimated\_length=unknown  
misc\_feature 1280. 2449  
/note="assembly\_name:Contig34"  
gap 2450. 2549  
/estimated\_length=unknown  
misc\_feature 2550. 4084  
/note="assembly\_name:Contig36"  
gap 4085. 4184  
/estimated\_length=unknown  
misc\_feature 4185. 6466  
/note="assembly\_name:Contig37"  
gap 6467. 6566

misc\_feature  
6567. 8700  
/note="assembly\_name:Contig38"  
gap 8701. 8800  
/estimated\_length=unknown  
misc\_feature 8801. 24080  
/note="assembly\_name:Contig39"  
gap 24081. 24180  
/estimated\_length=unknown  
misc\_feature 24181. 54114  
/note="assembly\_name:Contig40"  
gap 54115. 54214  
/estimated\_length=unknown  
misc\_feature 54215. 87114  
/note="assembly\_name:Contig41"  
gap 87115. 87214  
/estimated\_length=unknown  
misc\_feature 87215. 194232  
/note="assembly\_name:Contig42"

ORIGIN

Query Match 1.4%; Score 54; DB 14; Length 194232;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2281 ATGAAGAAGGCTGTGGCCACCTTATCCCTTCATGGAAAAAGAAAGAGAA 2334  
|||||  
DB 140285 ATGAAGAAGGCTGTGGCCACCTTATCCCTTCATGGAAAAAGAAAGAGAA 140232

RESULT 46

AC137163

LOCUS

DEFINITION

AC137163

unordered pieces.

AC137163

AC137163.1 GI:25073033

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Rattus;

1 (bases 1 to 228824)

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anylebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebrugge, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensu, L., Loulae, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,





repeat\_region complement(18214..18306)  
/rpt\_family="ID\_B1"

Query Match 1.4%; Score 54; DB.9: Length 236536;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2281 ATGAAGAAGCGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAA 2334  
|||||  
Db 53174 ATGAAGAAGCGCTGTTGGCCACCTTATCCCTTTTCATGGAAAAAGAAAGAA 53227  
|||||

RESULT 48  
AC113729  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-97A15, WORKING DRAFT SEQUENCE, 2  
unordered pieces.  
AC113729  
ACCESSION  
VERSION  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
SOURCE  
Rattus norvegicus (Norway rat)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.  
REFERENCE  
Muzny D, Marie L, Metzker M, Lee J, Abranzon S, Adams C, Alder J, Allen C, Allen H, Alsbrooks S, Amin A, Anguiano D, Anyalebech V, Aoyagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Bissal K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant M, Buhay C, Burch P, Burrell K, Calderon E, Cardenas V, Carter K, Cavazos I, Cesaar H, Center A, Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L, Davila M, Davis C, Davy-Carroll L, De Anda C, Dederich D, Delgado O, Denson S, Deramo C, Ding Y, Dinh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Fraser C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M, Gebrgeorgis E, Geer K, Gill R, Grady M, Guerra M, Guevara W, Gunaratne P, Haaland W, Hamill C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, Hodgson A, Hognes M, Hollins B, Howells S, Hulyk S, Hume J, Idlebird D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A, Karpach S, Kelly S, Khan Z, King L, Kovar C, Kowis C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J, Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J, Lorenshewa L, Loulesed H, Lozado R, Lu X, Ma J, Maheshwari M, Mahindartne M, Mahmoud M, Malloy K, Mangum A, Mangum B, Mapua P, Martin K, Martin R, Martine E, Mathew S, McLeod M, McNeill T, Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Pfannkuch C, Piopper P, Polindexter A, Popovic D, Primus E, Pu L, Puzo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R, Riley B, Reilly M, Ren Y, Reuter M, Richards S, Riggs P, Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders W, Savery G, Scherer S, Scott G, Shatman S, Shen H, Shetty J, Shvartsbeyn A, Sisson I, Sitter C, Smajd D, Snead A, Sodergren E, Song X, Z, Sorelle R, Sosa J, Steimle M, Strong R, Sutton A, Svatek A, Tabor P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Usmani K, Valas R, Vera V, Villaseca D, Waldron L, Walker B, Wang J, Wang Q, Wang S, Warren J, Warren R, Wei X, White F, Williams G, Willson R, Wleczek R, Wooden H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu F, Zhang J, Zhou J, Zhou X, Zhao S, Dunn D, von

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GSDI  
Center clone name: CH230-97A15  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 213541 bases at least Q40  
Consensus quality: 215589 bases at least Q30  
Consensus quality: 219247 bases at least Q20  
Estimated insert size: 221068; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 235052: contig of 235052 bp in length  
\* 235053 235152: gap of unknown length  
\* 235153 244504: contig of 9352 bp in length.  
----- Location/Qualifiers  
1. 244504  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-97A15"  
/notes="clone boundary  
144269..145175  
site\_end:Sp6  
end sequence:BH279190"  
235053..235152  
/estimated\_length-unknown  
235153..236174  
/notes="wgs end extension"

FEATURES  
source  
misc\_feature  
gap  
misc\_feature



```

ORIGIN
Query Match 1.4% Score 54; DB 14; Length 244504;
Best Local Similarity 100.0%; Pred. No. 1.2e-18;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2281 ATGAAGAGGCTGTGGCCACCTTATCCCTTCCTGATGGAAGAAAGAGAAGAA 2334
Db 180918 ATGAAGAGGCTGTGGCCACCTTATCCCTTCCTGATGGAAGAAAGAGAAGAA 180971

RESULT 49
AC121488/c
LOCUS AC121488
DEFINITION Rattus norvegicus clone CH230-unknown, WORKING DRAFT SEQUENCE, 6
unordered pieces.
ACCESSION AC121488
VERSION AC121488.3 GI:24941249
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
DeLaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlssoon, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kurehi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H.,
Lozago, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G.,
Oragunye, N., Owiedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 268715)
Worley, K.C.
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 268715)  
Worley, K.C.  
Direct Submission  
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 13, 2002 this sequence version replaced gi:23664539.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: TUVU  
Center clone name: CH230-unknown  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 247503 bases at least Q40  
Consensus quality: 250434 bases at least Q30  
Consensus quality: 252681 bases at least Q20  
Estimated insert size: 259015; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 6 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 3413: contig of 3413 bp in length  
3414 3513: gap of unknown length  
3514 16165: contig of 12652 bp in length  
16166 16265: gap of unknown length  
16266 94984: contig of 78719 bp in length  
94985 95084: gap of unknown length  
95085 266072: contig of 170988 bp in length  
266073 266172: gap of unknown length  
266173 267497: contig of 1325 bp in length  
267498 267597: gap of unknown length  
267598 268715: contig of 1118 bp in length.  
267598 Location/Qualifiers

FEATURES  
source

1. .268715  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-unknown"

misc\_feature 1. .1273  
/note="wgs contig"

gap 3414. .3513  
/estimated\_length=unknown

misc\_feature 3514. .4599  
/note="wgs contig"

gap 16166. .16265  
/estimated\_length=unknown

misc feature 16266. .17954

misc\_feature /note="wgs contig"  
19442..22918  
gap /note="wgs contig"  
94985..95084  
misc\_feature /estimated\_length=unknown  
95085..97304  
gap /note="wgs contig"  
266073..266172  
gap /estimated\_length=unknown  
267498..267597  
/estimated\_length=unknown

ORIGIN  
Query Match 1.4%; Score 54; DB 14; Length 268715;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2281 ATGAAGAAGGCTGTGGCCACCTTATCCCTTTCATGAAAGAGAGAGAA 2334  
|||||  
Db 46925 ATGAAGAAGGCTGTGGCCACCTTATCCCTTTCATGAAAGAGAGAGAA 46872  
|||||

RESULT 50  
AJ720857 5168 bp mRNA linear VRT 12-JAN-2005  
LOCUS  
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 27p10.  
ACCESSION AJ720857  
VERSION AJ720857.1 GI:53136373  
KEYWORDS ORF1.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.

REFERENCE 1  
AUTHORS Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J., Fiedler P., Kutter S., Blagodatki A., Kostovska D., Koter M., Plachy J., Carninci P., Hayashizaki Y. and Buerstedde J.M.  
TITLE Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis  
JOURNAL Genome Biol. 6 (1), R6 (2005)  
PUBMED 15642098  
REFERENCE 2 (bases 1 to 5168)  
AUTHORS Caldwell R.B.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum, Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr. 1, D-85764 Neuherberg, GERMANY

FEATURES  
source Location/Qualifiers  
1..5168  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="CB"  
/db\_xref="taxon:9031"  
/clone="27p10"  
/cell\_type="lymphocyte"  
/tissue\_type="bursa"  
/clone\_lib="rikeni"  
/dev\_stage="2 weeks old"  
1..5168  
/locus\_tag="RCJMB04\_27p10"  
41..31  
/locus\_tag="RCJMB04\_27p10"  
32..3793  
/locus\_tag="RCJMB04\_27p10"  
/note="ORF1"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAG32516.1"  
/db\_xref="GI:53136374"  
/db\_xref="UniProt/TREMBL:Q521C7"  
/translation="MARVADPPADQTESVEDEIESVLRERIMILDGQGTMIQQHAL  
SEEDFRGHEPKDHSKPLKGNLDSITQPDICDIKHKEYLLAGADIITETNTFSSTRA

gene  
5'UTR  
CDS

Query Match 1.0%; Score 41; DB 5; Length 5168;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2468 GAGTTATTGATTAGGAGTCATGCTCCATGTGATAGATA 2508  
|||||  
Db 2427 GAGTTATTGATTAGGAGTCATGCTCCATGTGATAGATA 2467  
|||||

Search completed: March 7, 2006, 07:08:45  
Job time : 18971 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 01:41:09 ; Search time 2063 Seconds  
(without alignments)  
12660.655 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcaccctgtgagagcagc.....ctcaagggaatacaacacctag 3919

Scoring table:

OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 18

Total number of hits satisfying chosen parameters: 42522

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002as.\*

7: Geneseqn2002bs.\*

8: Geneseqn2003as.\*

9: Geneseqn2003bs.\*

10: Geneseqn2003cs.\*

11: Geneseqn2003ds.\*

12: Geneseqn2004as.\*

13: Geneseqn2004bs.\*

14: Geneseqn2005s.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	89.7	7122	4	Aaf31100 Methionin
2	3517	89.7	7122	5	Aac91207 Human met
3	3517	89.7	7122	14	Aea63658 Methionin
4	3466	88.4	7224	4	Aaf31101 Methionin
5	3390	86.5	3798	10	Adi39109 Homo sapi
6	1349	34.4	3920	2	Aav34063 Human met
7	308	7.9	4172	13	Adro7169 Full leng
8	222	5.7	305	2	Aav88164 EST clone
9	44	1.1	80	12	Adp49459 Oligonuc1
10	31	0.8	31	4	Aai29993 Human sin
11	31	0.8	31	4	Aai29989 Human sin
12	31	0.8	31	4	Aai29985 Human sin
13	31	0.8	31	4	Aai29988 Human sin
14	31	0.8	31	4	Aai29986 Human sin
15	31	0.8	31	4	Aai29992 Human sin
16	31	0.8	31	4	Aai29987 Human sin
17	31	0.8	31	4	Aai29990 Human sin
18	31	0.8	31	4	Aai29991 Human sin
19	28	0.7	28	2	Aav34083 Methionin

c

c	20	27	0.7	27	12	ADO57182	Human MTR
21	26	26	0.7	26	2	AAV34086	Methionin
22	26	26	0.7	26	2	AAV34073	Methionin
23	26	26	0.7	26	2	AAV34091	Methionin
24	26	26	0.7	26	2	AAV34095	Methionin
25	26	26	0.7	26	2	AAV34101	Methionin
26	26	26	0.7	26	2	AAV34101	Methionin
27	25	25	0.6	25	2	AAV34097	Methionin
28	25	25	0.6	25	2	AAV34099	Methionin
29	25	25	0.6	25	2	AAV34098	Methionin
30	25	25	0.6	25	2	AAV34100	Methionin
31	25	25	0.6	25	2	AAV34096	Methionin
32	25	25	0.6	25	2	AAV34103	Methionin
33	25	25	0.6	25	2	AAV34087	Methionin
34	25	25	0.6	25	2	AAV34077	Methionin
35	25	25	0.6	25	2	AAV34079	Methionin
36	25	25	0.6	25	2	AAV34079	Methionin
37	25	25	0.6	25	2	ABL86554	Human ova
38	25	25	0.6	25	2	ACN52851	Cotton an
39	25	25	0.6	25	2	ACN52668	Cotton an
40	25	25	0.6	25	2	AAH03248	Human con
41	25	25	0.6	25	2	AAA61270	Human sec
42	25	25	0.6	25	2	ABX10897	DNA encod
43	25	25	0.6	25	2	ADK66235	Human NBO
44	25	25	0.6	25	2	ABA19751	Human ner
45	25	25	0.6	25	2	ABD32612	Human can
46	25	25	0.6	25	2	ABE96535	Human CAB
47	25	25	0.6	25	2	ADL13693	Osteocarth
48	25	25	0.6	25	2	AAV34076	Methionin
49	24	24	0.6	24	2	AAV34078	Methionin
50	24	24	0.6	24	2	AAV34088	Methionin
51	24	24	0.6	24	2	AAV34092	Methionin
52	24	24	0.6	24	2	AAV34094	Methionin
53	24	24	0.6	24	2	AAV34074	Methionin
54	24	24	0.6	24	2	AAV34090	Methionin
55	24	24	0.6	24	2	AAV34075	Methionin
56	24	24	0.6	24	2	AAV34104	Methionin
57	24	24	0.6	24	2	AAV34080	Methionin
58	24	24	0.6	24	2	AAV34102	Methionin
59	24	24	0.6	24	2	ADO57142	Human MTR
60	24	24	0.6	24	2	ADO57140	Human MTR
61	24	24	0.6	24	2	ACL54432	Human col
62	24	24	0.6	24	2	ABL86206	Human ova
63	24	24	0.6	24	2	ABL87802	Human ova
64	24	24	0.6	24	2	ABV54279	Human pro
65	24	24	0.6	24	2	ABQ59238	Human col
66	24	24	0.6	24	2	ABN87783	Human pro
67	24	24	0.6	24	2	AAI97716	Human neu
68	24	24	0.6	24	2	ADY17373	DNA encod
69	24	24	0.6	24	2	AAV36849	Human XLI
70	24	24	0.6	24	2	ABX76234	Lung canc
71	24	24	0.6	24	2	ADV70098	Tumor-ase
72	24	24	0.6	24	2	AAI99368	Human exc
73	24	24	0.6	24	2	AAI63718	Human kid
74	24	24	0.6	24	2	ABE96528	Human AP1
75	24	24	0.6	24	2	ADP65634	Human seq
76	24	24	0.6	24	2	ACN44956	Mouse gen
77	23	23	0.6	23	2	AAV34093	Methionin
78	23	23	0.6	23	2	AAV34084	Methionin
79	23	23	0.6	23	2	AAH49106	Human MTR
80	23	23	0.6	23	2	ABX09548	Arteriosc
81	23	23	0.6	23	2	ADO57141	Human MTR
82	23	23	0.6	23	2	ABT40755	Toxicity
83	23	23	0.6	23	2	ADP71801	Renal tox
84	23	23	0.6	23	2	ADV39794	Rat cardi
85	23	23	0.6	23	2	ABK62924	Rat seque
86	23	23	0.6	23	2	ADW21652	Rat seque
87	23	23	0.6	23	2	ABN73227	Bovine em
88	23	23	0.6	23	2	ABN73317	Bovine em
89	23	23	0.6	23	2	ABV56430	Human pol
90	23	23	0.6	23	2	AAI89372	Human pol
91	23	23	0.6	23	2	ADA44975	Human pol
92	23	23	0.6	23	2	AAV99731	Human adu

93	23	0.6	498	10	AD555663	Ad55663 Toxicity-	C 166	22	0.6	1031	10	ADI57327	ADI57327 Human GAB
94	23	0.6	498	13	ADV39607	Adv39607 Rat cardi	C 167	22	0.6	1082	13	ADR20714	Adr20714 Human oes
95	23	0.6	527	10	AD55532	Ad55532 Toxicity-	C 168	22	0.6	1113	3	AA51375	AA51375 D. melano
96	23	0.6	1458	10	AD563990	Ad563990 Human gen	C 169	22	0.6	1124	4	AA533253	AA533253 DNA encod
97	23	0.6	1458	10	AD563994	Ad563994 Human gen	C 170	22	0.6	1136	10	ADB61368	Adb61368 Rheumatol
98	23	0.6	1492	14	ADW11288	Adw11288 Human C-t	C 171	22	0.6	1211	3	AACT6729	AACT6729 Human ORF
99	23	0.6	1566	6	ABL41621	Ab141621 Human C-t	C 172	22	0.6	1215	13	ADR83094	Adr83094 Human sca
100	23	0.6	1739	10	ADC98698	Adc98698 Drosophil	C 173	22	0.6	1277	8	ABX05419	Abx05419 Human nov
101	23	0.6	1875	4	AAH33156	Aah33156 Human col	C 174	22	0.6	1349	12	ADP04768	Adp04768 Sea squir
102	23	0.6	1926	12	ADP07658	Adp07658 Human sec	C 175	22	0.6	1352	2	AAV41918	AAV41918 Nucleotid
103	23	0.6	2012	6	ABQ81083	Abq81083 Human dyn	C 176	22	0.6	1471	10	ADK66919	Adk66919 Gene #9 f
104	23	0.6	2100	13	ACN42458	Acn42458 Human dia	C 177	22	0.6	1471	13	ADR28922	Adr28922 Human Hee
105	23	0.6	2552	13	ACN38335	Acn38335 Tumour-as	C 178	22	0.6	1471	13	ADR83406	Adr83406 Human hai
106	23	0.6	3382	12	AD133246	Ad133246 Rat calci	C 179	22	0.6	1471	14	ADK70426	Adk70426 Human tra
107	23	0.6	4840	12	ADQ23344	Adq23344 Human sof	C 180	22	0.6	1518	2	AAK90449	AAK90449 Human sec
108	23	0.6	5565	6	ABL32263	Ab132263 Human imm	C 181	22	0.6	1518	4	AA59278	AA59278 Human cdn
109	23	0.6	8847	12	ADQ25304	Ado25304 Porcine u	C 182	22	0.6	1518	6	ABA90947	AbA90947 Human pol
110	23	0.6	9456	12	ADQ23327	Adq23327 Human sof	C 183	22	0.6	1518	14	ADM09230	Adm09230 cDNA clon
111	23	0.6	10480	6	ABL34201	Ab134201 Human imm	C 184	22	0.6	1520	2	AAZ24901	Aaz24901 Human sec
112	23	0.6	13288	5	ABA20468	AbA20468 Human ner	C 185	22	0.6	1538	9	ADA45144	Ada45144 Human pol
113	23	0.6	14117	6	ABL64107	Ab164107 Breast ca	C 186	22	0.6	1573	2	AAK35557	AAK35557 Secreted
114	23	0.6	17431	4	AAK90339	Aak90339 Human dig	C 187	22	0.6	1748	14	AEA36233	Aea36233 Human nuc
115	23	0.6	17431	4	AAI57710	Aai57710 Human col	C 188	22	0.6	1791	13	ADX61621	Adx61621 Plant ful
116	23	0.6	17431	6	ABS99887	Abs99887 Genomic D	C 189	22	0.6	1809	12	ADQ23540	Adq23540 Human sof
117	23	0.6	17431	10	ADB93040	Adb93040 Human col	C 190	22	0.6	1809	12	ADQ22903	Adq22903 Human sof
118	23	0.6	20486	6	ABL34611	Ab134611 Human met	C 191	22	0.6	1877	3	AACT79705	Aac79705 Human sec
119	23	0.6	20486	7	ADS99872	AdS99872 Complemen	C 192	22	0.6	1999	6	ABL62728	Ab162728 Colon ade
120	23	0.6	23271	13	ACN37213	Acn37213 Human per	C 193	22	0.6	2274	4	ABL22784	Ab122784 Drosophil
121	23	0.6	34245	13	ABD32614	Abd32614 Mouse can	C 194	22	0.6	2351	11	ADM03016	Adm03016 Human col
122	23	0.6	98825	14	AD245062	Continuation (10 o	C 195	22	0.6	2595	4	AAH33385	Aah33385 Human col
123	23	0.6	169659	12	ADQ59434	Adq59434 Human can	C 196	22	0.6	2828	14	AEA62873	Aea62873 Rat trans
124	23	0.6	187851	14	ADZ13735	Adz13735 Human can	C 197	22	0.6	3033	14	ADK98503	Adk98503 Human hyp
125	23	0.6	195102	13	ADV70461	Adv70461 Chicken g	C 198	22	0.6	3054	12	ADN00753	Adn00753 Human ROR
126	22	0.6	22	2	AAV34085	Aav34085 Methionin	C 199	22	0.6	3115	3	ACX98113	Acx98113 Human col
127	22	0.6	22	4	AAH49105	Aah49105 Human MTR	C 200	22	0.6	3157	8	ACA48896	ACA48896 Prokaryot
128	22	0.6	22	6	ABX09547	Abx09547 Arteriosc	C 201	22	0.6	3452	12	ADQ24824	Adq24824 Human sof
129	22	0.6	121	12	ADK93800	Adk93800 Polynucle	C 202	22	0.6	4182	5	ADL62107	Adl62107 Human ova
130	22	0.6	194	6	ABL86939	Ab186939 Human ova	C 203	22	0.6	4256	4	ABL24674	Ab124674 Drosophil
131	22	0.6	287	6	ABL85961	Ab185961 Human ova	C 204	22	0.6	4602	3	AA51404	AA51404 Chromosom
132	22	0.6	324	6	ABK62734	Abk62734 Rat seque	C 205	22	0.6	5413	4	AA46694	AA46694 Tumour su
133	22	0.6	324	10	AD511132	Ad511132 Primary r	C 206	22	0.6	6045	4	AAH18244	Aah18244 Human cdn
134	22	0.6	352	13	ACR95082	AcR95082 Human SIR	C 207	22	0.6	6450	2	AAZ23433	Aaz23433 Human est
135	22	0.6	374	6	ABV99252	Abv99252 Marine sn	C 208	22	0.6	6450	6	ABK89695	Abk89695 Oestrogen
136	22	0.6	375	12	ADP72025	Adp72025 Renal tox	C 209	22	0.6	6450	6	ABK99703	Abk99703 Oestrogen
137	22	0.6	375	14	ADZ5799	Adz5799 Novel cel	C 210	22	0.6	6450	6	ABV94431	Abv94431 Breast ca
138	22	0.6	424	5	ABV13705	Abv13705 Human pro	C 211	22	0.6	6450	6	ABZ23389	Abz23389 Nucleotid
139	22	0.6	436	5	ABV34818	Abv34818 Human pro	C 212	22	0.6	6450	6	ABZ23388	Abz23388 Nucleotid
140	22	0.6	458	4	AAH69080	Aah69080 Human cer	C 213	22	0.6	6450	8	ABX12085	Abx12085 cDNA enco
141	22	0.6	477	13	ACF83141	Acf83141 Human SIR	C 214	22	0.6	6450	8	ACC50140	Acc50140 Breast ca
142	22	0.6	482	4	AAI94479	Aai94479 Human neu	C 215	22	0.6	6450	8	ABX93797	Abx93797 Human oes
143	22	0.6	485	13	ACF83837	Acf83837 Human SIR	C 216	22	0.6	6450	10	AD881383	Ad881383 DNA seque
144	22	0.6	488	9	ACH37251	Ach37251 Human end	C 217	22	0.6	6450	10	ADD25499	Add25499 Binding d
145	22	0.6	490	4	AAH71943	Aah71943 Human cer	C 218	22	0.6	6450	10	ADE12135	Adel2135 Human oes
146	22	0.6	490	4	AAH72697	Aah72697 Human cer	C 219	22	0.6	6450	10	ADF76402	Adf76402 Novel hum
147	22	0.6	493	9	ACH33954	Ach33954 Human end	C 220	22	0.6	6450	10	ADG89355	Adg89355 Cancer de
148	22	0.6	506	5	ADL33740	Adl33740 Human ova	C 221	22	0.6	6450	12	ADG83093	Adg83093 Human tes
149	22	0.6	518	6	AAH62708	Aah62708 Colon ade	C 222	22	0.6	6450	12	ADF72329	Adf72329 Human and
150	22	0.6	550	4	AAH70743	Aah70743 Human cer	C 223	22	0.6	6450	12	ADP07306	Adp07306 Human ESR
151	22	0.6	560	10	ABT41123	Abt41123 Toxicity	C 224	22	0.6	6450	12	ADP05662	Adp05662 Human nuc
152	22	0.6	569	6	ABZ34987	Abz34987 Human gen	C 225	22	0.6	6450	13	ADR24574	Adr24574 Breast ca
153	22	0.6	569	10	ADE84813	AdE84813 Farnesyl	C 226	22	0.6	6450	13	ADR45377	Adr45377 Human oes
154	22	0.6	583	4	AAH08387	Aah08387 Human cdn	C 227	22	0.6	6450	13	ADR66804	Adr66804 Human pro
155	22	0.6	588	3	AAA16270	Aaa16270 Human col	C 228	22	0.6	6450	13	ADR65901	Adr65901 Human pro
156	22	0.6	598	5	ABV56139	Abv56139 Human pro	C 229	22	0.6	6450	14	ADW86780	Adw86780 Human ESR
157	22	0.6	669	1	AAH90103	Aah90103 Human pre	C 230	22	0.6	6450	14	ADY61188	Ady61188 Breast ca
158	22	0.6	688	11	ADT95691	Adt95691 Colon can	C 231	22	0.6	6450	14	ADY09607	Ady09607 Human bre
159	22	0.6	688	11	ADX42173	Adx42173 Human cdn	C 232	22	0.6	6450	14	AE25746	Aeb25746 Human tes
160	22	0.6	705	4	AAI24879	Aai24879 Human bre	C 233	22	0.6	6619	12	ADQ24081	Adq24081 Human sof
161	22	0.6	759	6	ABK53355	Abk53355 Human col	C 234	22	0.6	7171	6	ABN80216	Abn80216 Human che
162	22	0.6	812	10	ABE85970	AbE85970 Methyl-Cp	C 235	22	0.6	7510	6	ABT04094	Abt04094 Human imm
163	22	0.6	816	10	ABT42015	Abt42015 Toxicity	C 236	22	0.6	8042	6	ABL32927	Ab132927 Human imm
164	22	0.6	876	13	ADX59421	Adx59421 Plant ful	C 237	22	0.6	8168	6	ABL32928	Ab132928 Human imm
165	22	0.6	1028	6	ABK88911	Abk88911 Human agg	C 238	22	0.6	8168	6	AA563328	AA563328 Chemicall

C 239	22	0.6	8480	2	AAZ27521	Aaz27521 Interleuk	312	22	0.6	180557	6	ABN85750	Abn85750 Human BAC
C 240	22	0.6	8763	4	AAK89468	Aak89468 Human dig	313	22	0.6	185371	6	ABT10718	Abt10718 Human bre
C 241	22	0.6	9158	2	AAK30447	Aak30447 Human IL-	314	22	0.6	187851	14	ADZ13735	Adz13735 Human can
C 242	22	0.6	9799	5	AAF97860	Aaf97860 Human neu	315	22	0.6	191150	12	ADM69029	Adm69029 Human pia
C 243	22	0.6	10957	6	ABL33111	AbL33111 Human imm	C 316	22	0.6	191150	12	ADM69029	Adm69029 Human pia
C 244	22	0.6	11036	4	AAS45411	Aas45411 Chemcall	317	22	0.6	245531	13	ABD33022	Abd33022 Human can
C 245	22	0.6	11036	6	ABK28264	Abk28264 DNA trans	318	22	0.6	251364	13	ADV35003	Adv35003 Murine CD
C 246	22	0.6	12405	6	AAS45330	Aas45330 Chemcall	319	22	0.6	251364	13	ADV35003	Adv35003 Murine CD
C 247	22	0.6	12405	6	ABK28169	Abk28169 DNA trans	320	22	0.6	251364	13	ADV35003	Adv35003 Murine CD
C 248	22	0.6	12405	6	AAS61143	Aas61143 Human gen	C 321	22	0.6	260160	12	ADQ20017	Adq20017 Human sof
C 249	22	0.6	12774	4	AAK72738	Aak72738 Human imm	322	22	0.6	349989	10	ADC86916	Adc86916 Human GPC
C 250	22	0.6	12774	5	ABA15302	Abal5302 Human ner	C 323	21	0.5	21	2	AAV34089	Aav34089 Methionin
C 251	22	0.6	13182	10	ABX15028	Abx15028 Human gen	C 324	21	0.5	21	2	AAV34081	Aav34081 Methionin
C 252	22	0.6	13788	5	ABA19999	Abal9999 Human ner	C 325	21	0.5	21	2	AAV34082	Aav34082 Methionin
C 253	22	0.6	17142	4	AAS45498	Aas45498 Chemcall	326	21	0.5	21	4	AAF96203	Aaf96203 Human gen
C 254	22	0.6	17142	6	ABL34107	AbL34107 Human imm	327	21	0.5	51	4	AAZ27835	Aaz27835 Human SNP
C 255	22	0.6	17142	6	ABK28430	Abk28430 DNA trans	C 328	21	0.5	53	2	AAV29894	Aav29894 Target se
C 256	22	0.6	17200	4	AAAL37025	Aal37025 Human mus	C 329	21	0.5	121	12	ADK92447	Adk92447 Polynucle
C 257	22	0.6	17200	8	ABX60013	Abx60013 cDNA enco	C 330	21	0.5	141	5	AD168641	Adi68641 Human ova
C 258	22	0.6	17200	12	ADJ30763	Adj30763 Human mus	C 331	21	0.5	141	5	AD175004	Adi75004 Human ova
C 259	22	0.6	17528	6	ABL32601	AbL32601 Human imm	332	21	0.5	200	5	AD175008	Adi75008 Human ova
C 260	22	0.6	19139	4	AAK70995	Aak70995 Human imm	333	21	0.5	200	5	AD168645	Adi68645 Human ova
C 261	22	0.6	19338	14	AEA61189	Aea61189 Human DNA	C 334	21	0.5	285	4	AAK72170	Aak72170 Human imm
C 262	22	0.6	22475	12	ADQ97998	Adq97998 Human can	C 335	21	0.5	290	5	ABV19408	Abv19408 Human pro
C 263	22	0.6	22478	9	ADA02849	Ada02849 Mouse Cbx	336	21	0.5	292	5	ABV07826	Abv07826 Human pro
C 264	22	0.6	22478	10	ADB72587	Adb72587 Mouse Cbx	C 337	21	0.5	301	13	ACF90242	Acf90242 Human SIR
C 265	22	0.6	22478	10	ADC85328	Adc85328 Human Pgf	338	21	0.5	310	13	ADV40197	Adv40197 Rat cardli
C 266	22	0.6	22478	12	ADM74444	Adm74444 Murine ca	339	21	0.5	315	8	ABX47201	Abx47201 Bovine ES
C 267	22	0.6	22609	4	AAAL35781	Aal35781 Human mus	340	21	0.5	327	8	ABX47722	Abx47722 Bovine ES
C 268	22	0.6	22609	8	ABX58769	Abx58769 cDNA enco	341	21	0.5	337	5	ABV19636	Abv19636 Human pro
C 269	22	0.6	22609	12	ADJ29519	Adj29519 Human mus	C 342	21	0.5	339	5	ABV19636	Abv19636 Human pro
C 270	22	0.6	24707	6	ABQ74933	Abq74933 Human tra	343	21	0.5	355	6	ABU83749	Abu83749 Human ova
C 271	22	0.6	24707	9	ADA24263	Ada24263 Human tra	C 344	21	0.5	360	4	AAAL36865	Aal36865 Human mus
C 272	22	0.6	28444	6	ABK66948	Abk66948 Human glu	C 345	21	0.5	360	4	AAAL36865	Aal36865 Human mus
C 273	22	0.6	29764	14	ABE31548	Aeb31548 Glutathio	C 346	21	0.5	360	12	ADJ30603	Adj30603 Human mus
C 274	22	0.6	32193	4	AAAL35782	Aal35782 Human mus	347	21	0.5	371	6	ABU79958	Abu79958 Human ova
C 275	22	0.6	32193	8	ABX58770	Abx58770 cDNA enco	348	21	0.5	379	8	ABX37940	Abx37940 Bovine ES
C 276	22	0.6	32193	12	ADJ29520	Adj29520 Human mus	349	21	0.5	385	3	AAAC9443	Aac94443 Cat flea
C 277	22	0.6	32203	14	ADZ59521	Adz59521 Secondary	350	21	0.5	388	10	AAZ59584	Aaz59584 Human end
C 278	22	0.6	35000	6	ABT12817	Abt12817 Human REC	351	21	0.5	390	8	ABX44122	Abx44122 Bovine ES
C 279	22	0.6	36445	4	ABL18116	AbL18116 Drosophil	C 352	21	0.5	390	10	ADI62653	Adi62653 Human apo
C 280	22	0.6	38054	11	ACN44994	Acn44994 Human gen	353	21	0.5	392	4	AAI80366	Aai80366 Human pol
C 281	22	0.6	38054	11	ACN44994	Acn44994 Human gen	354	21	0.5	393	5	ABV13158	Abv13158 Human pro
C 282	22	0.6	41765	4	AAK76675	Aak76675 Human imm	355	21	0.5	393	5	ABV37751	Abv37751 Human pro
C 283	22	0.6	41772	4	AAK76676	Aak76676 Human imm	356	21	0.5	406	8	ABX54473	Abx54473 Bovine ES
C 284	22	0.6	41787	13	ABD33599	Abd33599 Human can	357	21	0.5	410	13	ADR60196	Adr60196 Cotton CD
C 285	22	0.6	42016	13	ABD32962	Abd32962 Human can	358	21	0.5	418	4	AAAL18463	Aal18463 Human bre
C 286	22	0.6	42334	9	ADA02588	Ada02588 Human ICS	C 359	21	0.5	418	13	ACN51983	Acn51983 Cotton an
C 287	22	0.6	42334	10	ADB72326	Adb72326 Human ICS	C 360	21	0.5	422	5	ABV48501	Abv48501 Human pro
C 288	22	0.6	42334	10	ADB72326	Adb72326 Human ICS	C 361	21	0.5	422	8	ABZ36470	Abz36470 Human GEN
C 289	22	0.6	43887	13	ABD33139	Abd33139 Murine ca	362	21	0.5	427	12	ADQ24433	Adq24433 Human sof
C 290	22	0.6	52691	4	AAK84435	Aak84435 Human imm	C 363	21	0.5	431	9	ACH21497	Ach21497 Human adu
C 291	22	0.6	59475	12	ADQ59530	Adq59530 Human can	364	21	0.5	437	6	ABQ55883	Abq55883 Human ova
C 292	22	0.6	61197	11	ACN45130	Acn45130 Human gen	C 365	21	0.5	442	5	ABV10111	Abv10111 Human pro
C 293	22	0.6	65237	6	ABA90193	AbA90193	C 366	21	0.5	442	13	ACN61251	Acn61251 Cotton gy
C 294	22	0.6	65237	6	ABA90193	AbA90193	C 367	21	0.5	444	5	ABV49403	Abv49403 Human pro
C 295	22	0.6	65237	8	ABX33717	AbX33717	C 368	21	0.5	444	5	ADL40245	Adl40245 Human ova
C 296	22	0.6	75252	11	ACN44450	Acn44450 Human gen	369	21	0.5	446	10	ADB55311	Adb55311 Toxicity-
C 297	22	0.6	76180	13	ABD33385	Abd33385 Human can	C 370	21	0.5	455	6	AAAL61528	Aal61528 Lung emal
C 298	22	0.6	79528	6	AAAL50814	Aal50814 Human can	371	21	0.5	457	4	AAAL53342	Aal53342 Human foe
C 299	22	0.6	79528	12	ADQ19802	Adq19802 Human sof	372	21	0.5	457	4	AAI32948	Aai32948 Probe #16
C 300	22	0.6	89210	14	ADL21391	AdL21391 Human can	373	21	0.5	457	4	ABA23120	AbA23120 Probe #15
C 301	22	0.6	91823	10	ADL13497	AdL13497 Osteoearth	374	21	0.5	457	4	AAK27051	Aak27051 Human bon
C 302	22	0.6	105413	12	ADI36512	Adi36512 Human kin	375	21	0.5	457	4	AAK01608	Aak01608 Human bra
C 303	22	0.6	110000	13	ABD32911	Abd32911	C 376	21	0.5	467	5	ABV00942	Abv00942 Human pro
C 304	22	0.6	110000	14	ADZ45062	AdZ45062	C 377	21	0.5	467	10	ADE59577	AdE59577 Rat gene
C 305	22	0.6	110000	14	ADZ45062	AdZ45062	C 378	21	0.5	472	4	AAK85582	Aak85582 Human imm
C 306	22	0.6	110096	6	ABN95044	Abn95044 Gene #154	C 379	21	0.5	472	5	ADL40241	Adl40241 Human ova
C 307	22	0.6	139573	10	ADH58564	Adh58564 Human Na+	380	21	0.5	472	11	ADT96089	Adt96089 Colon can
C 308	22	0.6	156416	13	ABD32817	Abd32817 Human can	381	21	0.5	472	11	ADX42571	Adx42571 Human cdn
C 309	22	0.6	169659	12	ADQ59434	Adq59434 Human can	C 382	21	0.5	478	4	AAK32968	Aak32968 Human bon
C 310	22	0.6	170245	12	ADP13586	Adp13586 Renal cel	C 383	21	0.5	478	4	AAK07213	Aak07213 Human bra
C 311	22	0.6	178024	12	ADQ97721	Adq97721 Human can	C 384	21	0.5	478	4	ABS32697	AbS32697 Human liv

C 385	21	0.5	478	6	ABS07776	AbS07776 Human gen	C 458	21	0.5	998	13	ADT16529	Adt16529 Plant cDN
C 386	21	0.5	482	11	ACN79991	Acn79991 Breast ca	C 459	21	0.5	999	10	ADK41704	Adk41704 Maize ami
C 387	21	0.5	485	6	ABV99324	Abv99324 Marine en	C 460	21	0.5	1108	4	AAF33113	Aaf33113 Human sec
C 388	21	0.5	486	5	ABV40252	Abv40252 Human pro	C 461	21	0.5	1114	10	ADF82206	Adf82206 Leukaemia
C 389	21	0.5	489	5	ABV31283	Abv31283 Human pro	C 462	21	0.5	1120	4	AAK56591	Aak56591 Human imm
C 390	21	0.5	486	5	ABA21253	AbA21253 Human ner	C 463	21	0.5	1120	4	AAK78749	Aak78749 Human imm
C 391	21	0.5	491	5	ABV40823	Abv40823 Human pro	C 464	21	0.5	1214	12	ADQ25563	Adm94267 Rice GTP
C 392	21	0.5	501	5	ABV20622	Abv20622 Human pro	C 465	21	0.5	1239	12	ADM94267	Adm94267 Rice GTP
C 393	21	0.5	501	5	ABV26463	Abv26463 Human pro	C 466	21	0.5	1243	8	ABX63096	Abx63096 Human cDN
C 394	21	0.5	511	5	ABV34278	Abv34278 Human pro	C 467	21	0.5	1260	4	AAK71520	Aak71520 Human imm
C 395	21	0.5	511	5	ABV43139	Abv43139 Human pro	C 468	21	0.5	1260	4	AAK71521	Aak71521 Human imm
C 396	21	0.5	524	5	ADL40167	Adl40167 Human ova	C 469	21	0.5	1260	4	AAK71522	Aak71522 Human imm
C 397	21	0.5	525	5	ABV55993	Abv55993 Human pro	C 470	21	0.5	1266	13	ACN39840	Acn39840 Tumour-as
C 398	21	0.5	525	13	ACN48293	Acn48293 Cotton pr	C 471	21	0.5	1292	6	ABQ93317	Abq93317 Human cDN
C 399	21	0.5	529	10	AAK82665	Aak82665 Rice endo	C 472	21	0.5	1312	10	ADI02590	Adi02590 Human cDN
C 400	21	0.5	540	4	AAK63674	Aak63674 Human imm	C 473	21	0.5	1318	10	ADB37417	Adb37417 Human can
C 401	21	0.5	559	4	AAH70901	Aah70901 Human cer	C 474	21	0.5	1318	13	ADR25077	Adr25077 Breast ca
C 402	21	0.5	565	13	ADQ57734	Adq57734 Novel can	C 475	21	0.5	1318	13	ADR14390	Adr14390 Human NF-
C 403	21	0.5	585	5	ADI74930	Adi74930 Human ova	C 476	21	0.5	1318	13	ACN40999	Acn40999 Tumour-as
C 404	21	0.5	585	5	ADI68566	Adi68566 Human ova	C 477	21	0.5	1318	13	ADP24031	Adp24031 PRO polyP
C 405	21	0.5	587	5	ABV22867	Abv22867 Human pro	C 478	21	0.5	1318	14	ADY15813	Ady15813 DNA encod
C 406	21	0.5	587	5	ABV28697	Abv28697 Human pro	C 479	21	0.5	1318	14	ADY61801	Ady61801 Human gen
C 407	21	0.5	597	5	AAH87676	Aah87676 Peppermin	C 480	21	0.5	1342	3	AAAC98822	Aac98822 Human pan
C 408	21	0.5	601	13	ADV15822	Adv15822 Human ost	C 481	21	0.5	1356	11	ABD15290	Abd15290 Pseudomon
C 409	21	0.5	601	13	ADV15365	Adv15365 Human ost	C 482	21	0.5	1393	14	ADY18383	Ady18383 DNA encod
C 410	21	0.5	601	13	ADV15366	Adv15366 Human ost	C 483	21	0.5	1413	6	ABQ76765	Abq76765 Human L1
C 411	21	0.5	601	13	ADV15821	Adv15821 Human ost	C 484	21	0.5	1496	4	AAH15218	Aah15218 Human cDN
C 412	21	0.5	601	13	ADV15367	Adv15367 Human ost	C 485	21	0.5	1521	10	AAU55798	Aau55798 CDNA enco
C 413	21	0.5	601	13	ADV15823	Adv15823 Human ost	C 486	21	0.5	1521	13	ADR25589	Adr25589 Breast ca
C 414	21	0.5	606	4	AAH09208	Aah09208 Human cDN	C 487	21	0.5	1521	13	ADP55051	Adp55051 Human PRO
C 415	21	0.5	636	11	ACN86945	Acn86945 Breast ca	C 488	21	0.5	1559	3	AAA26377	Aaa26377 Human sec
C 416	21	0.5	645	3	AAA75620	Aaa75620 Nucleotid	C 489	21	0.5	1559	12	ADL71438	Adl71438 Novel hum
C 417	21	0.5	645	6	AAAS20735	Aas20735 Human zal	C 490	21	0.5	1569	5	AAAS84119	Aas84119 DNA encod
C 418	21	0.5	645	10	ADH44678	Adh44678 Human zal	C 491	21	0.5	1580	6	ABQ54264	Abq54264 Human ova
C 419	21	0.5	645	10	ADIO1014	Adi01014 Human zal	C 492	21	0.5	1583	6	ABS61424	Abs61424 Prostate
C 420	21	0.5	645	12	ADP19837	Adp19837 Human zal	C 493	21	0.5	1616	6	ABL61093	AbL61093 Proteosom
C 421	21	0.5	645	14	ADV96461	Adv96461 Human zal	C 494	21	0.5	1639	5	AAH64730	Aah64730 Human sec
C 422	21	0.5	674	4	AAFI7977	Aaf17977 Human bre	C 495	21	0.5	1645	13	ADT17160	Adt17160 Plant cDN
C 423	21	0.5	674	4	AAI67220	Aai67220 B726p spl	C 496	21	0.5	1674	12	ADJ62777	Adj62777 Human cDN
C 424	21	0.5	674	4	AAAS47407	Aas47407 Human cDN	C 497	21	0.5	1674	13	ACN40026	Acn40026 Tumour-as
C 425	21	0.5	674	6	ABSE64008	Abse64008 Human bre	C 498	21	0.5	1681	4	AAF75080	Aaf75080 Human col
C 426	21	0.5	674	10	ABT33220	Abt33220 Human tum	C 499	21	0.5	1762	8	ABZ18595	Abz18595 Group III
C 427	21	0.5	674	12	ADL93127	Adl93127 Human bre	C 500	21	0.5	1778	4	AAAS42859	Aas42859 Human G P
C 428	21	0.5	674	11	ADSE44417	Adse44417 Human cDN	C 501	21	0.5	1791	10	ADP81426	Adp81426 Leukaemia
C 429	21	0.5	678	12	ADQ20094	Adq20094 Human sof	C 502	21	0.5	1803	12	ADQ24149	Adq24149 Human sof
C 430	21	0.5	679	3	AAFI14688	Aafi14688 Aspergill	C 503	21	0.5	1826	12	ADQ64818	Adq64818 Novel hum
C 431	21	0.5	679	13	ADU58729	Adu58729 Aspergill	C 504	21	0.5	1827	3	AAAC93322	Aac93322 Human sec
C 432	21	0.5	679	14	ADZ96732	Adz96732 Aspergill	C 505	21	0.5	1827	8	ABZ73648	Abz73648 Secreted
C 433	21	0.5	699	4	AAI84827	Aai84827 Human pol	C 506	21	0.5	1827	8	ADA98140	Ada98140 Human sec
C 434	21	0.5	733	13	ADP55241	Adp55241 Human PRO	C 507	21	0.5	1827	8	ADA43996	Ada43996 Human sec
C 435	21	0.5	735	8	ACF34515	Acf34515 Gene enco	C 508	21	0.5	1827	10	ADC20292	Adc20292 Human sec
C 436	21	0.5	772	3	AAFI15058	Aafi15058 Trichoder	C 509	21	0.5	1827	10	ADF10667	Adf10667 Human sec
C 437	21	0.5	772	13	ADU59099	Adu59099 Trichoder	C 510	21	0.5	1827	10	ABT16848	Abt16848 Human sec
C 438	21	0.5	772	14	ADZ97102	Adz97102 Trichoder	C 511	21	0.5	1827	10	ABZ67242	Abz67242 Human sec
C 439	21	0.5	806	4	AAI97414	Aai97414 Human neu	C 512	21	0.5	1829	4	AAH33888	Aah33888 Human col
C 440	21	0.5	806	4	AAH03234	Aah03234 Human cDN	C 513	21	0.5	1853	13	ADT18784	Adt18784 Plant cDN
C 441	21	0.5	807	8	ABT43046	Abt43046 Human neu	C 514	21	0.5	1861	10	ADL08456	Adl08456 Human can
C 442	21	0.5	838	14	ADX26034	Adx26034 Novel cel	C 515	21	0.5	1861	10	ADL08458	Adl08458 Human can
C 443	21	0.5	896	4	AAH33122	Aah33122 Human col	C 516	21	0.5	1868	5	ABV22237	Abv22237 Human pro
C 444	21	0.5	908	5	ABV07580	Abv07580 Human pro	C 517	21	0.5	1868	5	ABV28075	Abv28075 Human pro
C 445	21	0.5	913	4	AAAS3162	Aas3162 Human car	C 518	21	0.5	1918	4	AAF72744	Aaf72744 Human pro
C 446	21	0.5	913	5	ABA20711	AbA20711 Human ner	C 519	21	0.5	1927	4	AAAF72768	Aaf72768 Human pro
C 447	21	0.5	913	10	ADE46856	Ade46856 Human car	C 520	21	0.5	2000	7	ADZ74577	Adz74577 Arabidops
C 448	21	0.5	913	13	ADJ08274	Adj08274 Human car	C 521	21	0.5	2023	5	ABV24752	Abv24752 Human pro
C 449	21	0.5	914	4	AAAS36161	Aas36161 Human car	C 522	21	0.5	2056	4	AAAS01496	Aas01496 Human sec
C 450	21	0.5	914	5	ABA20710	AbA20710 Human ner	C 523	21	0.5	2056	4	AAH34935	Aah34935 Human col
C 451	21	0.5	914	10	ADE46855	Ade46855 Human car	C 524	21	0.5	2100	5	AAF24172	Aaf24172 Human sec
C 452	21	0.5	914	13	ADJ08273	Adj08273 Human car	C 525	21	0.5	2166	8	AAAD08390	Aad08390 Human sec
C 453	21	0.5	918	6	ABL90120	AbL90120 Human pol	C 526	21	0.5	2166	8	ABZ73348	Abz73348 Secreted
C 454	21	0.5	923	12	ADP07687	Adp07687 Human sec	C 527	21	0.5	2166	8	ADA97953	Ada97953 Human sec
C 455	21	0.5	931	4	AAAS32682	Aas32682 Human gen	C 528	21	0.5	2166	8	ADA43859	Ada43859 Human sec
C 456	21	0.5	931	4	AAAS32683	Aas32683 Human gen	C 529	21	0.5	2166	10	ADC20109	Adc20109 Human sec
C 457	21	0.5	932	4	AAK84143	Aak84143 Human imm	C 530	21	0.5	2166	10	ADF10572	Adf10572 Human sec



C 531	21	0.5	2177	13	ACN39039	Acn39039 Tumour-as	604	21	0.5	5470	4	AAS26619	Aas26619 Human gen
C 532	21	0.5	2222	5	ADM19255	Admi19255 Novel hum	605	21	0.5	5470	8	ABX73968	Abx73968 Human nov
C 533	21	0.5	2222	5	ADM19255	Admi19255 Novel hum	C 606	21	0.5	5559	6	ABL33423	Ab133423 Human imm
C 534	21	0.5	2233	5	ABV23302	Abv23302 Human pro	C 607	21	0.5	5605	13	ADS89678	Ads89678 Oligonucl
C 535	21	0.5	2233	5	ABV29149	Abv29149 Human pro	C 608	21	0.5	5605	13	ADS89404	Ads89404 Oligonucl
C 536	21	0.5	2249	14	AEA61198	Aea61198 Human FLTJ	C 609	21	0.5	5698	12	ADQ22854	Adq22854 Human sof
C 537	21	0.5	2249	14	AEA61198	Aea61198 Human FLTJ	C 610	21	0.5	5734	4	AAK67406	Aak67406 Human imm
C 538	21	0.5	2265	10	ACC00053	Acc00053 Hydroxyme	C 611	21	0.5	6019	6	ABK31191	Abk31191 Signal tr
C 539	21	0.5	2265	10	ADD05245	Add05245 Glutaryl	C 612	21	0.5	6019	6	ABL70160	Ab170160 Chemicall
C 540	21	0.5	2334	4	AAH72822	Aah72822 Human cer	C 613	21	0.5	6019	6	AAS61104	Aas61104 Human gen
C 541	21	0.5	2334	4	AAH72950	Aah72950 Human cer	C 614	21	0.5	6065	4	AAK65406	Aak65406 Human imm
C 542	21	0.5	2343	10	ADG10448	Adg10448 Human STA	C 615	21	0.5	6101	6	ABL33475	Ab133475 Human imm
C 543	21	0.5	2350	4	AAH17973	Aah17973 Human CDN	C 616	21	0.5	6127	6	ABL34449	Ab134449 Human met
C 544	21	0.5	2465	10	ADB63124	Adb63124 Human CDN	C 617	21	0.5	6127	6	ABL70120	Ab170120 Chemicall
C 545	21	0.5	2515	10	AAL56514	Aal56514 cDNA sequ	C 618	21	0.5	6127	7	ADS99710	Ads99710 Complemen
C 546	21	0.5	2519	10	ADD22468	Add22468 HLA-B46 T	C 619	21	0.5	6134	6	ABQ67140	Abq67140 Human ang
C 547	21	0.5	2519	10	AD115984	Ad115984 Human pp	C 620	21	0.5	6143	2	AAV42348	Aav42348 DNA encod
C 548	21	0.5	2566	5	ABA19833	Abal19833 Human ner	C 621	21	0.5	6167	6	ABK31407	Abk31407 Signal tr
C 549	21	0.5	2600	13	ADW67774	Adw67774 Human sec	C 622	21	0.5	6167	6	ABL70368	Ab170368 Chemicall
C 550	21	0.5	2657	12	ADQ25396	Adq25396 Human sof	C 623	21	0.5	6167	6	AAS61330	Aas61330 Human gen
C 551	21	0.5	2731	10	ADB62130	Adb62130 Human CDN	C 624	21	0.5	6283	6	ABL32089	Ab132089 Human imm
C 552	21	0.5	2749	11	ACN90411	Acn90411 Breast ca	C 625	21	0.5	6403	6	ABL33986	Ab133986 Human imm
C 553	21	0.5	2765	2	AAZ28810	Aaz28810 Rat membr	C 626	21	0.5	6431	6	ABQ67120	Abq67120 Human ang
C 554	21	0.5	2766	2	AAZ28810	Aaz28810 Rat membr	C 627	21	0.5	6432	4	AAQ65558	Aaq65558 Tumour su
C 555	21	0.5	2766	3	AAAT75488	Aaat75488 B. stearo	C 628	21	0.5	6432	10	ADB54290	Adb54290 Pretreate
C 556	21	0.5	2771	4	AAK91225	Aak91225 Human dig	C 629	21	0.5	6432	10	ADB54162	Adb54162 Pretreate
C 557	21	0.5	2776	4	AAK91226	Aak91226 Human dig	C 630	21	0.5	6432	10	ADB84124	Adb84124 Human lym
C 558	21	0.5	2804	11	ACN92346	Acn92346 Breast ca	C 631	21	0.5	6432	10	ADB84200	Adb84200 Human lym
C 559	21	0.5	2809	13	ADR24479	Adr24479 Breast ca	C 632	21	0.5	6432	13	ADS89316	Ads89316 Oligonucl
C 560	21	0.5	2846	2	AAV20806	Aav20806 Homo sapi	C 633	21	0.5	6432	13	ADS89590	Ads89590 Oligonucl
C 561	21	0.5	2846	3	AAK91006	Aak91006 Human VEG	C 634	21	0.5	6455	5	ABAI17328	Abal17328 Human ner
C 562	21	0.5	2937	3	AAK59340	Aak59340 Human sec	C 635	21	0.5	6455	5	AAS34621	Aas34621 Human DNA
C 563	21	0.5	2938	10	ADC51454	Adc51454 Human mac	C 636	21	0.5	6480	8	ACF34507	Acf34507 Gene enco
C 564	21	0.5	3227	6	ABK35482	Abk35482 Human end	C 637	21	0.5	6480	10	ADP76507	Adp76507 Novel hum
C 565	21	0.5	3227	6	ABK35482	Abk35482 Human end	C 638	21	0.5	6480	12	ADH17148	Adh17148 Human cxc
C 566	21	0.5	3228	6	ABV76643	Abv76643 Human mic	C 639	21	0.5	6480	12	ADL82848	Adl82848 Human PRO
C 567	21	0.5	3352	12	ADQ22709	Adq22709 Human sof	C 640	21	0.5	6480	12	ADP07302	Adp07302 Human CCN
C 568	21	0.5	3473	8	ABX76408	Abx76408 Lung canc	C 641	21	0.5	6480	12	ADO20473	Ado20473 Human PRO
C 569	21	0.5	3473	11	ADN39037	Adn39037 Cancer/an	C 642	21	0.5	6480	12	ADO19115	Ado19115 Human PRO
C 570	21	0.5	3523	8	ABX76407	Abx76407 Lung canc	C 643	21	0.5	6480	12	ADO19117	Ado19117 Human PRO
C 571	21	0.5	3523	9	ACC85484	Acc85484 Human alp	C 644	21	0.5	6480	13	ADP54150	Adp54150 Human PRO
C 572	21	0.5	3523	11	ADN39035	Adn39035 Cancer/an	C 645	21	0.5	6480	13	ADP23073	Adp23073 PRO polyv
C 573	21	0.5	3705	8	ACA42279	Aca42279 Prokaryot	C 646	21	0.5	6480	14	ADY14344	Ady14344 DNA encod
C 574	21	0.5	3705	10	AD139093	Ad139093 Pseudomon	C 647	21	0.5	6480	14	ADY14346	Ady14346 DNA encod
C 575	21	0.5	3748	4	AAS36163	Aas36163 Human car	C 648	21	0.5	6901	2	AAT78853	Aat78853 Human lec
C 576	21	0.5	3748	5	ABA20712	Abal20712 Human ner	C 649	21	0.5	6901	8	ADA26529	Ada26529 Human LCA
C 577	21	0.5	3748	10	AD846857	Ad846857 Human car	C 650	21	0.5	6930	6	ABQ93533	Abq93533 Human Dis
C 578	21	0.5	3748	13	ADJ08275	Adj08275 Human car	C 651	21	0.5	6930	13	ADR47378	Adr47378 Human DIS
C 579	21	0.5	3762	11	ABD14844	Abd14844 Pseudomon	C 652	21	0.5	6963	6	ABL32978	Ab132978 Human imm
C 580	21	0.5	3773	4	AAS36160	Aas36160 Human car	C 653	21	0.5	7458	3	AAA70106	Aaa70106 Plasmodiu
C 581	21	0.5	3773	5	ABA20709	Abal20709 Human ner	C 654	21	0.5	7513	4	AAK72990	Aak72990 Human imm
C 582	21	0.5	3773	10	AD846854	Ad846854 Human car	C 655	21	0.5	7513	4	AAL36017	Aal36017 Human mus
C 583	21	0.5	3773	13	ADJ08272	Adj08272 Human car	C 656	21	0.5	7513	8	ABX59005	Abx59005 cDNA enco
C 584	21	0.5	3789	12	ADQ23361	Adq23361 Human sof	C 657	21	0.5	7513	12	ADJ29755	Adj29755 Human mus
C 585	21	0.5	3795	11	ABD15089	Abd15089 Pseudomon	C 658	21	0.5	7928	4	ABK42648	Abk42648 Genomic s
C 586	21	0.5	3873	5	ABA19128	Abal19128 Human ner	C 659	21	0.5	7928	9	ADB60804	Adb60804 Connectiv
C 587	21	0.5	3892	10	ADC27642	Adc27642 Human col	C 660	21	0.5	8149	6	ABL33151	Ab133151 Human imm
C 588	21	0.5	3952	12	ADO22639	Ado22639 Human sof	C 661	21	0.5	8330	12	ADM66512	Adm66512 Human fib
C 589	21	0.5	3954	10	AD139105	Adi39105 Caulobact	C 662	21	0.5	8340	4	ABL29146	Ab129146 Drosophil
C 590	21	0.5	3966	4	AAK90316	Aak90316 Human dig	C 663	21	0.5	8605	6	ABL32535	Ab132535 Human imm
C 591	21	0.5	3966	4	AAI57690	Aai57690 Human col	C 664	21	0.5	8648	6	ABK31355	Abk31355 Signal tr
C 592	21	0.5	3966	6	ABS99867	Abs99867 Genomic D	C 665	21	0.5	8648	6	ABL70574	Ab170574 Chemicall
C 593	21	0.5	3966	10	ADB93020	Adb93020 Human col	C 666	21	0.5	8648	6	AAS61257	Aas61257 Human gen
C 594	21	0.5	4206	12	ADQ17424	Adq17424 Human sof	C 667	21	0.5	8996	4	AAS45504	Aas45504 Chemicall
C 595	21	0.5	4512	4	ABA07296	Abal07296 Human pan	C 668	21	0.5	8996	4	ABK28436	Abk28436 DNA trans
C 596	21	0.5	4512	4	AAK99941	Aak99941 Human dig	C 669	21	0.5	9293	4	AAS45320	Aas45320 Chemicall
C 597	21	0.5	4576	2	AAT96834	Aat96834 Intfion 4	C 670	21	0.5	9293	6	ABK39973	Abk39973 Human che
C 598	21	0.5	4648	13	ADRO6982	Adro6982 Full leng	C 671	21	0.5	9293	6	ABK28159	Abk28159 DNA trans
C 599	21	0.5	4660	14	ADZ70902	Adz70902 Human mat	C 672	21	0.5	9556	4	AAK80290	Aak80290 Human imm
C 600	21	0.5	5013	4	AAK74879	Aak74879 Human imm	C 673	21	0.5	9556	5	ADM20197	Adm20197 Alternati
C 601	21	0.5	5317	6	ABL32608	Ab132608 Human imm	C 674	21	0.5	9969	4	AAL04217	Aal04217 Human rep
C 602	21	0.5	5389	5	AAF98711	Aaf98711 Human lat	C 675	21	0.5	10296	4	AAK74868	Aak74868 Human imm
C 603	21	0.5	5391	6	ABL32537	Ab132537 Human imm	C 676	21	0.5	10467	6	ABL49301	Ab149301 Human pol

c 677	21	0.5	10825	6	ABN80206	Abn80206 Human che	750	21	0.5	31168	10	ADB94607	ADB94607 Novel hum
c 678	21	0.5	10953	5	AAS29204	Aas29204 Genomic s	751	21	0.5	31277	11	ACN44854	Acn44854 Human gen
c 679	21	0.5	10953	6	ABS68344	Abs68344 Human DNA	752	21	0.5	31405	4	AAK83153	Aak83153 Human imm
c 680	21	0.5	10953	10	ADC25466	Adc25466 Human CDN	753	21	0.5	31405	4	AAK74865	Aak74865 Human imm
c 681	21	0.5	10989	6	ABD33870	Abd33870 Human tra	c 754	21	0.5	31405	4	AAK67293	Aak67293 Human imm
c 682	21	0.5	10989	8	ABX95979	Abx95979 Human tra	c 755	21	0.5	31814	10	AAAD47150	AAad47150 Human Rse
c 683	21	0.5	13076	13	ADS89377	Ads89377 Oligonucl	756	21	0.5	31842	9	ADA02966	Ada02966 Human LCK
c 684	21	0.5	13076	13	ADS89651	Ads89651 Oligonucl	757	21	0.5	31842	10	ADB72704	Adb72704 Human LCK
c 685	21	0.5	13337	4	ABK42742	Abk42742 Genomic s	758	21	0.5	31842	10	ADC85446	Adc85446 Human Lck
c 686	21	0.5	13337	5	AAS29203	Aas29203 Genomic s	759	21	0.5	31842	12	ADM74561	Adm74561 Human car
c 687	21	0.5	13337	5	ABA18582	Abal18582 Human ner	760	21	0.5	31934	4	AAK82215	Aak82215 Human imm
c 688	21	0.5	13337	6	ABS68343	Abs68343 Human DNA	c 761	21	0.5	31934	4	AAS30619	Aas30619 DNA encod
c 689	21	0.5	13337	9	ADB60898	Adb60898 Connectiv	c 762	21	0.5	31994	4	AAS28165	Aas28165 Genomic s
c 690	21	0.5	13337	10	ADC25465	Adc25465 Human CDN	c 763	21	0.5	31994	8	ACA03382	AcA03382 DNA encod
c 691	21	0.5	13418	4	AAK86473	Aak86473 Human imm	c 764	21	0.5	31994	9	AB96730	Ab96730 Novel lun
c 692	21	0.5	13574	6	ABL33316	Ab133316 Human imm	c 765	21	0.5	31994	10	ADG41361	Adg41361 Human res
c 693	21	0.5	14537	4	AAS46356	Aas46356 Tumour su	c 766	21	0.5	31994	11	ADI97135	Adi97135 Human res
c 694	21	0.5	15122	6	ABL32775	Ab132775 Human imm	c 767	21	0.5	32148	4	AAL04218	Aal04218 Human rep
c 695	21	0.5	15224	6	ABK31358	Abk31358 Signal tr	c 768	21	0.5	32148	5	AAS30115	Aas30115 Human lun
c 696	21	0.5	15224	6	ABL70317	Ab170317 Chemicali	c 769	21	0.5	32189	10	ADB33452	Adb33452 Human nov
c 697	21	0.5	15224	6	AAS61262	Aas61262 Human gen	c 770	21	0.5	32221	5	AAS30113	Aas30113 Human lun
c 698	21	0.5	15914	4	AAK84889	Aak84889 Human imm	c 771	21	0.5	32221	10	ADB33450	Adb33450 Human nov
c 699	21	0.5	15914	8	ABZ73768	Abz73768 Secreted	c 772	21	0.5	33317	11	ACN44990	Acn44990 Human gen
c 700	21	0.5	15914	8	ADA98447	Ada98447 Human sec	c 773	21	0.5	34739	11	ACN45078	Acn45078 Human gen
c 701	21	0.5	15914	8	ADA44234	Ada44234 Human sec	c 774	21	0.5	34766	14	ADZ13291	Adz13291 Human can
c 702	21	0.5	15914	10	ADC20616	Adc20616 Human sec	c 775	21	0.5	35962	8	ABZ09958	Abz09958 Haematopo
c 703	21	0.5	15914	10	ADT10823	Adt10823 Human sec	c 776	21	0.5	35962	8	ABZ10104	Abz10104 Haematopo
c 704	21	0.5	15914	10	ABT16886	Abt16886 Human sec	c 777	21	0.5	36785	4	AAK82208	Aak82208 Human imm
c 705	21	0.5	15914	10	ABZ67365	Abz67365 Human sec	c 778	21	0.5	37590	4	AAS12439	Aas12439 DNA encod
c 706	21	0.5	16736	4	ABL26478	Ab126478 Drosophi	c 779	21	0.5	37790	12	ADK41243	Adk41243 DNA regio
c 707	21	0.5	17717	4	AAK82007	Aak82007 Human imm	c 780	21	0.5	38166	12	ADK41244	Adk41244 DNA regio
c 708	21	0.5	17717	4	AAK82008	Aak82008 Human imm	c 781	21	0.5	43058	6	ABL64982	Ab164982 Lung canc
c 709	21	0.5	17719	4	AAK82006	Aak82006 Human imm	c 782	21	0.5	43058	6	ABL65219	Ab165219 Lung canc
c 710	21	0.5	17874	4	AAK71070	Aak71070 Human imm	c 783	21	0.5	43058	6	ABN97455	Abn97455 Gene #395
c 711	21	0.5	18389	4	AAK82632	Aak82632 Human imm	c 784	21	0.5	49561	4	AAK82012	Aak82012 Human imm
c 712	21	0.5	18389	4	AAK74864	Aak74864 Human imm	785	21	0.5	49806	11	ACN44528	Acn44528 Mouse gen
c 713	21	0.5	19183	4	AAK64938	Aak64938 Human imm	786	21	0.5	50849	6	ABN87883	Abn87883 Human glu
c 714	21	0.5	19183	4	AAK36434	Aal36434 Human mus	c 787	21	0.5	51001	12	ADJ10262	Adj10262 Human ger
c 715	21	0.5	19183	8	ABX59422	Abx59422 cDNA enco	788	21	0.5	52302	9	ADA02738	Ada02738 Human CCN
c 716	21	0.5	19183	12	ADJ30172	Adj30172 Human mus	789	21	0.5	52302	10	ADB72476	Adb72476 Human CCN
c 717	21	0.5	19701	4	AAK91227	Aak91227 Human dig	c 790	21	0.5	52302	10	ADC85218	Adc85218 Human CCN
c 718	21	0.5	19814	6	ABL70609	Ab170609 Chemicali	c 791	21	0.5	52302	12	ADM74333	Adm74333 Human car
c 719	21	0.5	20467	4	AAK36361	Aal36361 Human mus	c 792	21	0.5	53122	11	ACN43998	Acn43998 (7 of
c 720	21	0.5	20467	4	AAK36360	Aal36360 Human mus	c 793	21	0.5	55404	11	ACN43988	Acn43988 Mouse gen
c 721	21	0.5	20467	8	ABX59348	Abx59348 cDNA enco	794	21	0.5	58366	10	ADC87618	Adc87618 Human GPC
c 722	21	0.5	20467	8	ABX59349	Abx59349 cDNA enco	795	21	0.5	59446	10	ADU47504	Adu47504 Human tra
c 723	21	0.5	20467	12	ADJ30098	Adj30098 Human mus	796	21	0.5	62944	6	ABL68262	Ab168262 Kidney can
c 724	21	0.5	20467	12	ADJ30099	Adj30099 Human mus	797	21	0.5	62944	6	ABL66947	Ab166947 Lung canc
c 725	21	0.5	21000	13	ADT77194	Adt77194 Type II d	798	21	0.5	69648	12	ADQ97934	Adq97934 Human can
c 726	21	0.5	21354	4	AAS46815	Aas46815 Tumour su	c 799	21	0.5	77777	12	ADQ94674	Adq94674 Human chr
c 727	21	0.5	21537	6	ABL33998	Ab133998 Human imm	800	21	0.5	85920	14	ADZ13418	Adz13418 Human can
c 728	21	0.5	21602	13	ABD33092	Abd33092 Human can	801	21	0.5	93538	13	ABD33442	Abd33442 Murine ca
c 729	21	0.5	21890	4	ABL115986	Ab115986 Drosophi	c 802	21	0.5	93538	13	ABD33442	Abd33442 Murine ca
c 730	21	0.5	22313	4	ABL11964	Ab111964 Drosophi	c 803	21	0.5	95914	13	ABD33444	Abd33444 Human can
c 731	21	0.5	23241	5	AAF97871	Aaf97871 Human neu	804	21	0.5	96596	9	ADA02564	Ada02564 Human RAS
c 732	21	0.5	23241	5	AAF97870	Aaf97870 Human neu	805	21	0.5	96596	10	ADB72302	Adb72302 Human RAS
c 733	21	0.5	23695	6	ADQ66982	Adq66982 Human ang	806	21	0.5	96596	10	ADQ95812	Adq95812 Human RAS
c 734	21	0.5	24292	5	ABA15939	Abal15939 Human ner	c 807	21	0.5	98638	12	ADQ97919	Adq97919 Mouse can
c 735	21	0.5	25000	6	ABL68348	Ab168348 Kidney ca	c 808	21	0.5	98844	7	ADJ84099	Adj84099 Human WRN
c 736	21	0.5	26040	4	AAH27887	Aah27887 Nucleotid	809	21	0.5	99960	3	AAS20905	Aas20905 Human TBC
c 737	21	0.5	26354	12	ADQ97467	Adq97467 Human can	c 810	21	0.5	101169	12	ADQ97584	Adq97584 Mouse can
c 738	21	0.5	27411	11	ACN44474	Acn44474 Human gen	c 811	21	0.5	110000	5	AAI61373	AAi61373 3
c 739	21	0.5	27411	14	ADZ12987	Adz12987 Human can	812	21	0.5	110000	9	ADJ13411	Adj13411 2
c 740	21	0.5	28313	4	AAK36829	Aak36829 Human mus	813	21	0.5	110000	10	ADG70447	Adg70447 2
c 741	21	0.5	28313	8	ABX59817	Abx59817 cDNA enco	814	21	0.5	110000	10	ABH10017	Abh10017 2
c 742	21	0.5	28313	12	ADJ30567	Adj30567 Human mus	815	21	0.5	110000	10	ABZ79565	Abz79565 2
c 743	21	0.5	28470	4	AAK71145	Aak71145 Human imm	c 816	21	0.5	110000	10	ABQ84281	AbQ84281 2
c 744	21	0.5	28854	8	AAK95685	Aak95685 Human gen	c 817	21	0.5	110000	11	ACN44150	Acn44150 3
c 745	21	0.5	31168	4	AAK27804	Aak27804 DNA encod	818	21	0.5	110000	11	ACN43998	Acn43998 5
c 746	21	0.5	31168	4	ABA07295	AbA07295 Human pan	819	21	0.5	110000	11	ACN44934	Acn44934 0
c 747	21	0.5	31168	4	AAS42115	Aas42115 Genomic s	820	21	0.5	110000	12	ADQ59446	AdQ59446 2
c 748	21	0.5	31168	4	AAK89940	Aak89940 Human dig	821	21	0.5	110000	12	ADQ97266	AdQ97266 2
c 749	21	0.5	31168	4	AAK64760	Aak64760 Human imm	822	21	0.5	110000	12	ADQ97138	AdQ97138 1

C 823	21	0.5	110000	12	ADQ97138_2	Continuation (3 of	896	21	0.5	294575	14	AEA61217	Aea61217 Human STR
C 824	21	0.5	110000	14	ADZ13757_2	Continuation (3 of	C 897	21	0.5	294575	14	AEA61217	Aea61217 Human STR
C 825	21	0.5	110000	14	ADZ12821_2	Continuation (3 of	C 898	21	0.5	301477	13	ABD33362	Abd33362 Human can
C 826	21	0.5	110000	14	ADZ12617_0	Continuation (3 of	C 899	21	0.5	301477	13	ABD33362	Abd33362 Human can
C 827	21	0.5	110000	14	ADZ46976_2	Continuation (3 of	C 900	21	0.5	313287	13	ABD333100	Abd333100 Human can
C 828	21	0.5	110000	14	ADZ46976_2	Continuation (3 of	C 901	21	0.5	313287	13	ABD333100	Abd333100 Human can
C 829	21	0.5	112604	12	ADZ42274_0	Adz42274 Human min	C 902	21	0.5	321491	11	ACN44202	Acn44202 Human can
C 830	21	0.5	116624	2	AAV52850	Adq18153 Human sof	C 903	21	0.5	325791	4	AA543104	Aas43104 Human gen
C 831	21	0.5	117328	13	ABD33886	Abd33886 Mouse can	C 904	21	0.5	334462	10	ADC24763	Adc24763 Human will
C 832	21	0.5	119057	13	ABD33460	Abd33460 Murine ca	C 905	21	0.5	335199	10	ADC24763	Adc24763 Human will
C 833	21	0.5	119950	4	AAX90201	Aax90201 Human his	C 906	21	0.5	349881	10	ADC86642	Adc86642 Human GPC
C 834	21	0.5	122186	4	AAC89560	Aac89560 Human yes	C 907	20	0.5	21	2	AA075669	Aa075669 Reverse t
C 835	21	0.5	124987	12	ADP03055	Adp03055 Human hou	C 908	20	0.5	28	2	AA770114	Aat70114 PolyAB pr
C 836	21	0.5	124990	13	ADP03055	Adp03055 Human hou	C 909	20	0.5	28	2	AA770114	Aat70114 PolyAB pr
C 837	21	0.5	124990	13	ADP03055	Adp03055 Human hou	C 910	20	0.5	28	2	AA770114	Aat70114 PolyAB pr
C 838	21	0.5	126001	12	ADH77123	Adh77123 Human PAZ	C 911	20	0.5	30	13	ADR27781	Adi27781 MTR gene
C 839	21	0.5	127238	11	ACN44372	Acn44372 Mouse gen	C 912	20	0.5	30	13	ADR27781	Adi27781 MTR gene
C 840	21	0.5	127238	11	ACN44372	Acn44372 Mouse gen	C 913	20	0.5	30	13	ADR27781	Adi27781 MTR gene
C 841	21	0.5	127272	12	ADQ97301	Adq97301 Human can	C 914	20	0.5	31	13	ADV95692	Adv95692 Hepatitis
C 842	21	0.5	127767	13	ABD33584	Abd33584 Murine ca	C 915	20	0.5	31	13	ADV95692	Adv95692 Hepatitis
C 843	21	0.5	127767	13	ABD33584	Abd33584 Murine ca	C 916	20	0.5	31	13	ADV95692	Adv95692 Hepatitis
C 844	21	0.5	127767	13	ABD33584	Abd33584 Murine ca	C 917	20	0.5	31	13	ADV95692	Adv95692 Hepatitis
C 845	21	0.5	129042	11	ACN44674	Acn44674 Human gen	C 918	20	0.5	51	4	AA132453	Aal32453 Human SNP
C 846	21	0.5	133632	11	ACN45054	Acn45054 Human gen	C 919	20	0.5	51	4	AA132453	Aal32453 Human SNP
C 847	21	0.5	133642	14	ADZ13285	Adz13285 Human gen	C 920	20	0.5	132	6	ABL85667	Ab185667 Human ova
C 848	21	0.5	134738	11	ACN44182	Acn44182 Human gen	C 921	20	0.5	133	5	AD172651	Adi72651 Human ova
C 849	21	0.5	142299	10	ADV50651	Adv50651 BAC seque	C 922	20	0.5	133	5	AD172651	Adi72651 Human ova
C 850	21	0.5	142299	10	ADV77909	Adv77909 Human BAC	C 923	20	0.5	155	13	ACF91490	Acf91490 Human SIR
C 851	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 924	20	0.5	155	13	ACF91490	Acf91490 Human SIR
C 852	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 925	20	0.5	168	13	ADJ12811	Adj12811 DNA fragm
C 853	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 926	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 854	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 927	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 855	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 928	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 856	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 929	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 857	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 930	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 858	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 931	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 859	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 932	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 860	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 933	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 861	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 934	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 862	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 935	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 863	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 936	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 864	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 937	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 865	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 938	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 866	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 939	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 867	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 940	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 868	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 941	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 869	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 942	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 870	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 943	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 871	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 944	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 872	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 945	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 873	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 946	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 874	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 947	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 875	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 948	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 876	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 949	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 877	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 950	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 878	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 951	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 879	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 952	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 880	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 953	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 881	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 954	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 882	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 955	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 883	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 956	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 884	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 957	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 885	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 958	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 886	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 959	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 887	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 960	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 888	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 961	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 889	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 962	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 890	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 963	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 891	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 964	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 892	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 965	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 893	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 966	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 894	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 967	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm
C 895	21	0.5	143776	10	ADP67269	Adp67269 Human chr	C 968	20	0.5	190	12	ADJ12811	Adj12811 DNA fragm

c 969 20 0.5 301 4 AAK67724  
 c 970 20 0.5 301 4 AAK84088  
 c 971 20 0.5 301 12 ADJ12789  
 c 972 20 0.5 305 4 AAK85349  
 c 973 20 0.5 313 4 AAS38991  
 c 974 20 0.5 315 6 ABL85650  
 c 975 20 0.5 315 10 ADF90608  
 c 976 20 0.5 316 5 ABV58014  
 c 977 20 0.5 317 4 AAL11061  
 c 978 20 0.5 317 4 AAK89039  
 c 979 20 0.5 317 5 AAS39638  
 c 980 20 0.5 317 9 ADB32598  
 c 981 20 0.5 319 10 ADD33011  
 c 982 20 0.5 326 12 ADP66311  
 c 983 20 0.5 331 5 ABV54228  
 c 984 20 0.5 334 4 AAH69569  
 c 985 20 0.5 334 4 AAH73022  
 c 986 20 0.5 347 6 ABL85185  
 c 987 20 0.5 349 10 ADP89921  
 c 988 20 0.5 350 7 ADS72585  
 c 989 20 0.5 350 7 ADM41439  
 c 990 20 0.5 355 5 AAS69983  
 c 991 20 0.5 355 5 AAS80266  
 c 992 20 0.5 356 4 AAL10919  
 c 993 20 0.5 357 5 ABV07938  
 c 994 20 0.5 361 10 ADE60435  
 c 995 20 0.5 362 4 AAL24388  
 c 996 20 0.5 363 4 AAK63668  
 c 997 20 0.5 363 6 AAK53893  
 c 998 20 0.5 366 4 AAI84933  
 c 999 20 0.5 366 5 ABV47779  
 1000 20 0.5 366 5 AAF65945

## ALIGNMENTS

RESULT 1  
 ID AAF31100 standard; cDNA; 7122 BP.  
 AC AAF31100;  
 DT 27-APR-2001 (first entry)  
 XX Methionine synthase coding sequence #1.  
 XX Analyte-binding enzyme; analyte analysis; ss.  
 XX Homo sapiens.  
 OS WO200102600-A2.  
 PN 11-JAN-2001.  
 PD 30-JUN-2000; 2000WO-US018057.  
 PF 06-JUL-1999; 99US-00347878.  
 PR 06-DEC-1999; 99US-00457205.  
 XX (GEAT ) GEN ATOMICS.  
 XX Yuan C;  
 PI WPI; 2001-071583/08.  
 DR Assaying method, useful for prognosis and diagnosis of disease, comprises  
 PT contacting sample with a mutant analyte-binding enzyme and detecting  
 PT binding.  
 PS Disclosure; Page; 187pp; English.  
 XX The present invention relates to a method for assaying an analyte in a

Aak67724 Human imm  
 Aak84088 Human imm  
 Adj12789 DNA fragm  
 Aak85349 Human imm  
 Aas38991 Novel hum  
 Abl85650 Human ova  
 Adf90608 Human hep  
 Abv58014 Human pro  
 Aal11061 Human bre  
 Aak89039 Human dig  
 Aas39638 Genomic s  
 Adb32598 Human nov  
 Add33011 Human mit  
 Adp66311 Human cdn  
 Abv54228 Human pro  
 Aah69569 Human cer  
 Aah73022 Human cer  
 Abl85185 Human ova  
 Add89921 Murine DN  
 Adh72585 Human kid  
 Adw41439 cDNA elev  
 Aas69983 DNA encod  
 Aas80266 DNA encod  
 Aal10919 Human bre  
 Abv07938 Human pro  
 Ade60435 Rat gene  
 Aal24388 Human bre  
 Aak63668 Human imm  
 Abk53893 Human hea  
 Aai84933 Human pol  
 Abv47779 Human pro  
 Aaf65945 Novel hum

CC sample comprising: contacting the sample with a mutant analyte-binding  
 CC enzyme which has binding affinity for the analyte or an immediate analyte  
 CC enzymatic conversion product but has attenuated catalytic activity; and  
 CC detecting resulting binding. The method is useful in monitoring  
 CC biological systems/processes, or prognosis/diagnosis of disease caused by  
 CC imbalances of the analytes. The present sequence is a coding sequence  
 CC used in the present invention. Note: the present sequence is not shown in  
 CC the specification, but was from Genbank, using information given in the  
 CC specification  
 CC  
 XX  
 SQ Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;  
 Query Match 89.7%; Score 3517; DB 4; Length 7122;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 2 GTCACCTGTGGAGAGACGCTCTCTCTGCGCGGCCCTCTGCGCAAGAGGCTCTGGAAGAAACCCCTCGGG 121  
 DB 225 GTCACCTGTGGAGAGACGCTCTCTCTGCGCGGCCCTCTGCGCAAGAGGCTCTGGAAGAAACCCCTCGGG 284  
 QY 62 ACATGTGACCCCGGCTCCAAAGACCTGTGCAACCCGAGGCTCTGGAAGAAACCCCTCGGG 121  
 DB 285 ACATGTGACCCCGGCTCCAAAGACCTGTGCAACCCGAGGCTCTGGAAGAAACCCCTCGGG 344  
 QY 122 ATGAGATCAATGCAATCTCTGCAAGAGGATATGCTGCTGATGAGGATGCGGACCA 181  
 DB 345 ATGAGATCAATGCAATCTCTGCAAGAGGATATGCTGCTGATGAGGATGCGGACCA 404  
 QY 182 TGATCAGCGGGAGAGAGCTTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241  
 DB 405 TGATCAGCGGGAGAGAGCTTAAACGAAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 464  
 QY 242 CAGGCGGCTGAAGGCGCAACATGACATTTTAAATTAATCACTCAGCTGATGCTATTACC 301  
 DB 465 CAGGCGGCTGAAGGCGCAACATGACATTTTAAATTAATCACTCAGCTGATGCTATTACC 524  
 QY 302 AAATCATAGGAATACCTTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTTAGCA 361  
 DB 525 AAATCATAGGAATACCTTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTTAGCA 584  
 QY 362 GCATAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCTTACCGGATGAAATGT 421  
 DB 585 GCATAGTATTTGCCCAAGCTGACTATGGCTTTGAACACTTGGCTTACCGGATGAAATGT 644  
 QY 422 GCTCTGAGGAGTGGCGAGAGAGCTTCCGAGAGGCTTACTCTCCAGACAGGAATTAAGA 481  
 DB 645 GCTCTGAGGAGTGGCGAGAGAGCTTCCGAGAGGCTTACTCTCCAGACAGGAATTAAGA 704  
 QY 482 GGTCTGAGGAGGCTTGGCTCCGACTTAATAGACACTCTCTGTGTCCCACTCTGTGG 541  
 DB 705 GGTCTGAGGAGGCTTGGCTCCGACTTAATAGACACTCTCTGTGTCCCACTCTGTGG 764  
 QY 542 AAAGGCGGATATAGGAACATCACATTTTGAAGAGCTTGTGAAGCATACCAAGAGCAGG 601  
 DB 765 AAAGGCGGATATAGGAACATCACATTTTGAAGAGCTTGTGAAGCATACCAAGAGCAGG 824  
 QY 602 CCAAGGACTCTGGATGGCGGGTGGATATCTTACTCATTTGAAACTATTTTGTACTG 661  
 DB 825 CCAAGGACTCTGGATGGCGGGTGGATATCTTACTCATTTGAAACTATTTTGTACTG 884  
 QY 662 CCAATGCCAAGGAGGCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATCTCCCC 721  
 DB 885 CCAATGCCAAGGAGGCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATCTCCCC 944  
 QY 722 GGCTATCTTTTATTTCAAGGAGCATGTTGATAAAGTGGCGGACTCTTTTCCGAGCAGA 781  
 DB 945 GGCTATCTTTTATTTCAAGGAGCATGTTGATAAAGTGGCGGACTCTTTTCCGAGCAGA 1004  
 QY 782 CAGGAGGAGGATTTGTCATCAGCGGTGCTCATGAGAGAACCACTCTGCTGATTTGAATT 841  
 DB 1005 CAGGAGGAGGATTTGTCATCAGCGGTGCTCATGAGAGAACCACTCTGCTGATTTGAATT 1064  
 QY 842 GTGCTTTGGGTGCGAGCTGAGATGAGACCTTTTATTGAAATAATTTGGAATAATGTACAACAG 901

1065 GTCTTTGGTGCAGCTGAATGAGACCTTTTATTGAAATAATTGGAATAATTGAAACAG 1124  
1092 CCTATGCTCTCTGTTATCCCAATGCGAGCTTCCCAACACCTTTGGTGTACTATGATGAAA 961  
1125 CCTATGCTCTCTGTTATCCCAATGCGAGCTTCCCAACACCTTTGGTGTACTATGATGAAA 1184  
962 CGCTTCTATGATGGCCAGCAGCTTAAAGAAATTTGTTATGATGAGTGGTGTGCTCAATATAG 1021  
1185 CGCTTCTATGATGGCCAGCAGCTTAAAGAAATTTGTTATGATGAGTGGTGTGCTCAATATAG 1244  
1022 TTGAGAGATGCTCTGGGTCAACACAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1081  
1245 TTGAGAGATGCTCTGGGTCAACACAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1304  
1082 ATTGTAAGCCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGTGGTC 1141  
1305 ATTGTAAGCCTTAGAGTTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGTGGTC 1364  
1142 TAGAGCCTTTCAGGATTTGACCGGTACACCACTTTGTTAACTTTGAGAGGCGCTGTAATG 1201  
1365 TAGAGCCTTTCAGGATTTGACCGGTACACCACTTTGTTAACTTTGAGAGGCGCTGTAATG 1424  
1202 TTGAGAGATGCTCTGGGTCAACACAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1261  
1425 TTGAGAGATGCTCTGGGTCAACACAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1484  
1262 GTCTTCCCAAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCACATGGATGATG 1321  
1485 GTCTTCCCAAGTGCAGGTGGAATGGAGCCAGGTTGGATGTCACATGGATGATG 1544  
1322 GCATGCTAGATGTCAGGATTTGCACTGATGATGATGATGATGATGATGATGATGATGATG 1381  
1545 GCATGCTAGATGTCAGGATTTGCACTGATGATGATGATGATGATGATGATGATGATG 1604  
1382 ACATCCCAAGTACCTTTGTCATGCACTGCTCCCAATTTGCTGCTGATGATGATGATG 1441  
1605 ACATCCCAAGTACCTTTGTCATGCACTGCTCCCAATTTGCTGCTGATGATGATGATG 1664  
1442 TAAAGTCTGCCAAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1501  
1665 TAAAGTCTGCCAAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1724  
1502 ACTTCTTGAGAGAGCCAGGATTTAAAGTATGAGCTGCTATGCTGCTGCTGCTGCTGCT 1561  
1725 ACTTCTTGAGAGAGCCAGGATTTAAAGTATGAGCTGCTATGCTGCTGCTGCTGCTGCT 1784  
1562 TTGATGAGAGAGCAGGAGCAAGAGACAGACAGACAGACAGACAGACAGACAGACAG 1621  
1785 TTGATGAGAGAGCAGGAGCAAGAGACAGACAGACAGACAGACAGACAGACAGACAG 1844  
1622 ACCATCTGCTTGTGAAGAACTGGCTTTAATCAATGACATTAATTTTGAACCTTAATA 1681  
1845 ACCATCTGCTTGTGAAGAACTGGCTTTAATCAATGACATTAATTTTGAACCTTAATA 1904  
1682 TCCTAACTTGGAGCTGGAATGGAGAAACACACTTTGATGCTGCTGCTGCTGCTGCTGCT 1741  
1905 TCCTAACTTGGAGCTGGAATGGAGAAACACACTTTGATGCTGCTGCTGCTGCTGCTGCT 1964  
1742 CAACAAAGTCATTAAGAAACATTAACCTGGAGCCAGATTAAGTGAAGCTTTTCAACT 1801  
1965 CAACAAAGTCATTAAGAAACATTAACCTGGAGCCAGATTAAGTGAAGCTTTTCAACT 2024  
1802 TGTCTCTTCTCTCCGAGGAATGGAAGCCATTCGAGAGCAATGAGTGGGTTTCTCTTT 1861  
2025 TGTCTCTTCTCTCCGAGGAATGGAAGCCATTCGAGAGCAATGAGTGGGTTTCTCTTT 2084  
1862 ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGATGATGATGATGATGATGATG 1921  
2085 ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGATGATGATGATGATGATGATG 2144  
1922 ATGATGATATCCATAAGAACTTCTGAGCTCTGTGAGATCTCTGAGATTAAGAGAC 1981

2145 ATGATGATATCCATTAAGAACTTCTGACCTCTGTGAAGATCTCATCTGGAATAAGACC 2204  
1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2041  
2205 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGAGGAGAAAG 2264  
2042 TCATTCAGACTGATGATGAGAAATGGCCCTGTGGAAGAAAGCCCTTGATGATGCTCTTG 2101  
2265 TCATTCAGACTGATGATGAGAAATGGCCCTGTGGAAGAAAGCCCTTGATGATGCTCTTG 2324  
2102 TGAAGGGCAATTCAAAAACATATTTTGAAGTACTGAGGAAGCCAGGTTAAAAACAAAAA 2161  
2325 TGAAGGGCAATTCAAAAACATATTTTGAAGTACTGAGGAAGCCAGGTTAAAAACAAAAA 2384  
2162 AATATCCCGACCTCTCAATATAATTAAGAGACCCCTGTATGAATGAATGAATGAATGTTG 2221  
2385 AATATCCCGACCTCTCAATATAATTAAGAGACCCCTGTATGAATGAATGAATGAATGTTG 2444  
2222 GTGATCTTTTGGAGCTGGAAGAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2281  
2445 GTGATCTTTTGGAGCTGGAAGAAATGTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2504  
2282 TGAAGAGGCTCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAAAGAGAAACCCAGAG 2341  
2505 TGAAGAGGCTCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAAAGAGAAACCCAGAG 2564  
2342 TGCCTTAAAGGACAGTAGAAGAGAGGACCTTTACAGGGCACCATCGTGTGGCCACTG 2401  
2565 TGCCTTAAAGGACAGTAGAAGAGAGGACCTTTACAGGGCACCATCGTGTGGCCACTG 2624  
2402 TTAAGGCGAGCTGCACGACATAGGCAAGAAATAGTTGGAGTAGTCTTGGTGTGAATA 2461  
2625 TTAAGGCGAGCTGCACGACATAGGCAAGAAATAGTTGGAGTAGTCTTGGTGTGAATA 2684  
2462 ATTTCCGAGCTTATGATTTAGGAGTCACTGCTTCAATGATGAATGATGAAGCTGCTC 2521  
2685 ATTTCCGAGCTTATGATTTAGGAGTCACTGCTTCAATGATGAATGATGAAGCTGCTC 2744  
2522 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCACTCTCTCTCTGGATGAAA 2581  
2745 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCACTCTCTCTCTGGATGAAA 2804  
2582 TGATTTTGTGCGCAAGGAAATGAGAGATTAAGCTTATAAGGATTCATTTGATTTGGAG 2641  
2805 TGATTTTGTGCGCAAGGAAATGAGAGATTAAGCTTATAAGGATTCATTTGATTTGGAG 2864  
2642 GAGCAACACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACTG 2701  
2865 GAGCAACACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACTG 2924  
2702 TAATCCATGCTCTGGACCGCTCAAGAGTGTGGTGTGTTTCCAGCTGTTAGATGAAA 2761  
2925 TAATCCATGCTCTGGACCGCTCAAGAGTGTGGTGTGTTTCCAGCTGTTAGATGAAA 2984  
2762 ATCTAAGATGAATATCTTTGAGGAAAATCATGGAAGAAATATGAAGATTTAGACAGGACC 2821  
2985 ATCTAAGATGAATATCTTTGAGGAAAATCATGGAAGAAATATGAAGATTTAGACAGGACC 3044  
2822 ATTATGAGTCTCTCAAGGAGAGGATCTTACCTTAACTTAACTTAACTTAACTTAACTT 2881  
3045 ATTATGAGTCTCTCAAGGAGAGGATCTTACCTTAACTTAACTTAACTTAACTTAACTT 3104  
2882 TCCAAATGATTTGGCTGTCTGAACCTTCAACAGTGAAGCCACCTGTTTATGAGGACCCAGG 2941  
3105 TCCAAATGATTTGGCTGTCTGAACCTTCAACAGTGAAGCCACCTGTTTATGAGGACCCAGG 3164  
2942 TCTTTGAAGATATGATCTGCAAGAGCTGGTGGAGCTATCAATGATGATGATGATGATGATG 3001  
3165 TCTTTGAAGATATGATCTGCAAGAGCTGGTGGAGCTTACATGATGATGATGATGATGATG 3224  
3002 ATGCTGAGGAGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGACCA 3061  
3225 ATGCTGAGGAGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGACCA 3284

```
QY 3062 AAACAGTAGGTGAGAGGCCAGGAGGCTTACGATGATGCCACATATGCTGAACAC 3121
DB 3285 AAACAGTAGGTGAGAGGCCAGGAGGCTTACGATGATGCCACATATGCTGAACAC 3344
QY 3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGTGTGGTTGGTTCTGGCCAGACAGAGTA 3181
DB 3345 TGATTAGTCAAAAGAACTCCGGGCCCGGGTGTGGTTGGTTCTGGCCAGACAGAGTA 3404
QY 3182 TCCAAGACGACATTCACCTGTATCGCAGAGGCTGCTGTGCCCCAGGCTGACAGGCCCATAG 3241
DB 3405 TCCAAGACGACATTCACCTGTATCGCAGAGGCTGCTGTGCCCCAGGCTGACAGGCCCATAG 3464
QY 3242 CCACTTTCTATGGGTTAAGCAACAGGCTGAGAAAGACTCTGCCAGCAGGAGCCATCT 3301
DB 3465 CCACTTTCTATGGGTTAAGCAACAGGCTGAGAAAGACTCTGCCAGCAGGAGCCATCT 3524
QY 3302 ACTGCCCTCTCAGACTTCATCGCTCCCTTGTGATCTGCTGATCCCTGATCTACTGGGCTGT 3361
DB 3525 ACTGCCCTCTCAGACTTCATCGCTCCCTTGTGATCTGCTGATCCCTGATCTACTGGGCTGT 3584
QY 3362 TTGCCGTTGCTCTTTGGGGTGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAG 3421
DB 3585 TTGCCGTTGCTCTTTGGGGTGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAG 3644
QY 3422 ACTACAGCAGCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCCCTTTGCAGAG 3481
DB 3645 ACTACAGCAGCATGCTCAAGCGCTGGGGACCGGCTGGCAGAGGCCCTTTGCAGAG 3704
QY 3482 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGAGC 3541
DB 3705 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGAGC 3764
QY 3542 TCGCAGACCTGCGAAGGTTGGGTACAAAGGCGATCCGCCGGCTCTGGCTACCCAGCC 3601
DB 3765 TCGCAGACCTGCGAAGGTTGGGTACAAAGGCGATCCGCCGGCTCTGGCTACCCAGCC 3824
QY 3602 AGCCCGACCAACCGAGAGCTACCATGTGGAGACTCGCAGACATCGACAGTCTACAG 3661
DB 3825 AGCCCGACCAACCGAGAGCTACCATGTGGAGACTCGCAGACATCGACAGTCTACAG 3884
QY 3662 GCATTAGGTTAAACAGATATTAGCAATGGCAGCTGCTCAGCAGTCTCAGGCTCTACT 3721
DB 3885 GCATTAGGTTAAACAGATATTAGCAATGGCAGCTGCTCAGCAGTCTCAGGCTCTACT 3944
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 3781
DB 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 4004
QY 3782 ATTATGCAATTGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841
DB 4005 ATTATGCAATTGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4064
QY 3842 TGGGATATGATACAGACTTAACCTTTTTTTTTTTTTT 3875
DB 4065 TGGGATATGATACAGACTTAACCTTTTTTTTTTTTTT 4098
```

## RESULT 2

AAC91207

ID AAC91207 standard; DNA; 7122 BP.

XX AAC91207;

XX AAC91207;

DT 20-MAR-2001 (first entry)

XX Human methionine synthase gene SEQ ID NO: 2.

XX Human; schizophrenia; developmental disorder; spina bifida cystica;  
XX Tourette's syndrome; bipolar illness; autism; conduct disorder;  
XX attention deficit disorder; obsessive compulsive disorder;  
XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.

OS Homo sapiens.

XX WO200071754-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US014354.

XX 25-MAY-1999; 99US-00318448.

XX (UYNE-) UNIV NEW JERSEY MEDICINE &amp; DENTISTRY.

XX Johnson WG, Stenroos ES;

XX WPI; 2001-025174/03.

XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming  
PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)  
PT and environmental variables affecting an individual and then comparing  
PT these DS with reference DS.

XX Disclosure; Page 115-117; 156pp; English.

XX The present invention provides a novel method of estimating the  
CC susceptibility of an individual to a developmental disorder using genetic  
CC and environmental variables. The method can be used in the diagnosis,  
CC prevention and treatment of disorders such as schizophrenia, spina bifida  
CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,  
CC attention deficit disorder, obsessive compulsive disorder, chronic  
CC multiple tic syndrome and learning disorders such as dyslexia

XX Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;

XX Query Match 89.7%; Score 3517; DB 5; Length 7122;  
XX Best Local Similarity 99.8%; Pred. No. 0;

XX Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GTCACTGTGGAGACACGCTCTCTCTGGCGGCCCTCTGCGCAAGGAGGAGACTCGACA 61

DB 225 GTCACTGTGGAGACACGCTCTCTCTGGCGGCCCTCTGCGCAAGGAGGAGACTCGACA 284

QY 62 ACATGTCAACCCCGCTCCAGAGCTCTGCGCAACCCGAGGTCTGAAGAAACCCCTCGGG 121

DB 285 ACATGTCAACCCCGCTCCAGAGCTCTGCGCAACCCGAGGTCTGAAGAAACCCCTCGGG 344

QY 122 ATGAGATCAATGTCATCTTGCAAGAGGATTTGTTGCTGGATGGAGGATGGGACCA 181

DB 345 ATGAGATCAATGTCATCTTGCAAGAGGATTTGTTGCTGGATGGAGGATGGGACCA 404

QY 182 TGATCCAGCGGAGAGAGCTTAAACGAAAGACACCTTCGAGGTCAAGGATTTAAAGATCATG 241

DB 405 TGATCCAGCGGAGAGAGCTTAAACGAAAGACACCTTCGAGGTCAAGGATTTAAAGATCATG 464

QY 242 CCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTATACTCAGCCTGATGTCTATTACC 301

DB 465 CCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTATACTCAGCCTGATGTCTATTACC 524

QY 302 AAATCCATAAGGAATACTTCTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 361

DB 525 AAATCCATAAGGAATACTTCTGCTGGGCGAGATATCATTTGAAACAAATACTTTTAGCA 584

QY 362 GCACTAGTATTTGCCAGCTGACTATGGCCCTTGAACACTTTGGCTACCGGATGAACATGT 421

DB 585 GCACTAGTATTTGCCAGCTGACTATGGCCCTTGAACACTTTGGCTACCGGATGAACATGT 644

QY 422 GCTCTGAGGAGTGGCCAGAAAAGCTCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 481

DB 645 GCTCTGAGGAGTGGCCAGAAAAGCTCCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 704

QY 482 GGTGTTGGCAGGGGCTCTGGGTCGAGCTAATAAGACACTCTCTGTGTCCTCCCATCTGGG 541

DB 705 GGTGTTGGCAGGGGCTCTGGGTCGAGCTAATAAGACACTCTCTGTGTCCTCCCATCTGGG 764



QY 542 AAAGCGGATATTAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACG 601  
Db 765 AAAGCGGATATTAGGAACATCACATTTGATGAGCTTGTGAAGCATACCAAGACG 824  
QY 602 CCAAGGACTTCTGGATGGCGGGTGTGATCTTACTCATTTGAAACTATTTTGTGATACTG 661  
Db 825 CCAAGGACTTCTGGATGGCGGGTGTGATCTTACTCATTTGAAACTATTTTGTGATACTG 884  
QY 662 CCAATGCCAAGGACGCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
Db 885 CCAATGCCAAGGACGCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 944  
QY 722 GGCCTATCTTTATTTCAAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGGACAGA 781  
Db 945 GGCCTATCTTTATTTCAAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGGACAGA 1004  
QY 782 CAGGAGGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTCATTTGGATTAAAT 841  
Db 1005 CAGGAGGAGGATTTGTCTATCAGCGTGTCTCATGAGAACCACTCTCATTTGGATTAAAT 1064  
QY 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATAATTTGAAATGTACACAG 901  
Db 1065 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTTGAAATAATTTGAAATGTACACAG 1124  
QY 902 CCTATGTCCTCTGTATCCCAATGCAAGTCTTCCCAACACCTTTGCTGACTATGATGAA 961  
Db 1125 CCTATGTCCTCTGTATCCCAATGCAAGTCTTCCCAACACCTTTGCTGACTATGATGAA 1184  
QY 962 CGCCTTCTATGATGGGCAAGCACTTAAAGGATTTGCTATGATGCTTGGTCAATATAG 1021  
Db 1185 CGCCTTCTATGATGGGCAAGCACTTAAAGGATTTGCTATGATGCTTGGTCAATATAG 1244  
QY 1022 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1081  
Db 1245 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAGCTTAGATTTCCACTGCCACTGCTTTTGAAGACATATGTTACTGTCTGGTC 1141  
Db 1305 ATTGTAAGCTTAGATTTCCACTGCCACTGCTTTTGAAGACATATGTTACTGTCTGGTC 1364  
QY 1142 TAGAGCCCTTTCAGGATTTGAGACCTGACACCAACTTTTGAACATTTGAGAGCGCTGTAATG 1201  
Db 1365 TAGAGCCCTTTCAGGATTTGAGACCTGACACCAACTTTTGAACATTTGAGAGCGCTGTAATG 1424  
QY 1202 TTGACGATCAAGGAATTTGCTTAACTCATATGCGAGAAACTATGAAGAGCCCTTGT 1261  
Db 1425 TTGACGATCAAGGAATTTGCTTAACTCATATGCGAGAAACTATGAAGAGCCCTTGT 1484  
QY 1262 GTGTTGCCAAGTGCAAGTGGAAATGGAGCCCGAGGTGTGGATGTCAACATGGATGATG 1321  
Db 1485 GTGTTGCCAAGTGCAAGTGGAAATGGAGCCCGAGGTGTGGATGTCAACATGGATGATG 1544  
QY 1322 GCATGCTAGATGCTCAAGTGCAATGACACAGATTTTGAACCTTAAATGCTTCCGAGCCAG 1381  
Db 1545 GCATGCTAGATGCTCAAGTGCAATGACACAGATTTTGAACCTTAAATGCTTCCGAGCCAG 1604  
QY 1382 ACATCGCAAGGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1441  
Db 1605 ACATCGCAAGGTACCTTTGTGATCGACTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1664  
QY 1442 TAAAGTGTGCGCAAGGAGATGCTATGTCATAGCAATTTAGTCTGAAGGAGGAGAGAG 1501  
Db 1665 TAAAGTGTGCGCAAGGAGATGCTATGTCATAGCAATTTAGTCTGAAGGAGGAGAGAGAG 1724  
QY 1502 ACTTCTTGAGAGGCGCAGGAGATTTAAAGTATGAGCTGCTATGCTGCTCATGGCTT 1561  
Db 1725 ACTTCTTGAGAGGCGCAGGAGATTTAAAGTATGAGCTGCTATGCTGCTCATGGCTT 1784  
QY 1562 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATATCAGAGTGTGCACCCGGGCT 1621  
Db 1785 TTGATGAAGAGGACAGGCAACAGAAACAGACACAAATATCAGAGTGTGCACCCGGGCT 1844  
QY 1622 ACCATCTGCTTGTGAAAAAACTGGGCTTTAAATCCAAATGACATTAATTTTGGACCCCTAATA 1681

Db 1845 ACCATCTGCTTGTGAAAAAACTGGGCTTTAAATCCAAATGACATTAATTTTGGACCCCTAATA 1904  
QY 1682 TCCTAACCAATTTGGGACTGGAATGGAGGAACACAACCTTGTATGCCATTAATTTTATCCATG 1741  
Db 1905 TCCTAACCAATTTGGGACTGGAATGGAGGAACACAACCTTGTATGCCATTAATTTTATCCATG 1964  
QY 1742 CAAACAAAGTCATTAAGAAAAATTTACCTTGGAGCCAGAAATAGTGAGAGTCTTTTCCAACT 1801  
Db 1965 CAAACAAAGTCATTAAGAAAAATTTACCTTGGAGCCAGAAATAGTGAGAGTCTTTTCCAACT 2024  
QY 1802 TGTCTCTTCTCTTCCGAGGAATGGAGCCATTTCCGAGAACATATGCAATGCAATGGGTTTTTCCCTT 1861  
Db 2025 TGTCTCTTCTCTTCCGAGGAATGGAGCCATTTCCGAGAACATATGCAATGCAATGGGTTTTTCCCTT 2084  
QY 1862 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCCCTGTGT 1921  
Db 2085 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCCCTGTGT 2144  
QY 1922 ATGATGATATCCATTAAGAACTTCTGCACTCTGTGAAGATCTCATCTGGAAATAAGAC 1981  
Db 2145 ATGATGATATCCATTAAGAACTTCTGCACTCTGTGAAGATCTCATCTGGAAATAAGAC 2204  
QY 1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGSCACAGAGGGAAGAAAG 2041  
Db 2205 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGSCACAGAGGGAAGAAAG 2264  
QY 2042 TCATTCAGACTCATGATGGAGAAATGGCCCTGTCCGAGAACCTTGTAGTATGCCCTTG 2101  
Db 2265 TCATTCAGACTCATGATGGAGAAATGGCCCTGTCCGAGAACCTTGTAGTATGCCCTTG 2324  
QY 2102 TGAAGGGCAATGAAAAACATATTTTGAAGTATCTGAGAAAGCCAGTTTAAACCAAAAAA 2161  
Db 2325 TGAAGGGCAATGAAAAACATATTTTGAAGTATCTGAGAAAGCCAGTTTAAACCAAAAAA 2384  
QY 2162 AATNTCCCGACTCTCATATATAATTTGAAGGACCCCTGATGAATGGAATGAATTTGTTG 2221  
Db 2385 AATATCCCGACTCTCATATATAATTTGAAGGACCCCTGATGAATGGAATGAATTTGTTG 2444  
QY 2222 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGCTTATAAGTCAAGCCCGGTTA 2281  
Db 2445 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGCTTATAAGTCAAGCCCGGTTA 2504  
QY 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTTATGAAAAAGAAAGAGAAACAGAG 2341  
Db 2505 TGAAGAGGCTGTGGCCACCTTATCCCTTTTATGAAAAAGAAAGAGAAACAGAG 2564  
QY 2342 TGCCTTAAAGGCAAGTGAAGAAAGAGGACCCCTTACAGGGCCACCATGCTGCTGGCCACTG 2401  
Db 2565 TGCCTTAAAGGCAAGTGAAGAAAGAGGACCCCTTACAGGGCCACCATGCTGCTGGCCACTG 2624  
QY 2402 TTAAAGGGGACCTGCACGACATAGGCAAGAAACATAGTTGAGTGTGCTTGGCTGCAATA 2461  
Db 2625 TTAAAGGGGACCTGCACGACATAGGCAAGAAACATAGTTGAGTGTGCTTGGCTGCAATA 2684  
QY 2462 ATTTCCGAGTTATGATTTAGAGTCTGATCTTCAATGATGAAGTATGAAAGCTGCTC 2521  
Db 2685 ATTTCCGAGTTATGATTTAGAGTCTGATCTTCAATGATGAAGTATGAAAGCTGCTC 2744  
QY 2522 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTCTCTCTGGATGAAA 2581  
Db 2745 TTGACCAACAAAGCAGATATAATTTGGCTGTGAGGACTCATCTCTCTCTCTGGATGAAA 2804  
QY 2582 TGAATTTTGTCCCAAGGAATGGAGAGATTTAGCTATAGGATTCATTTGTTGATTGGAG 2641  
Db 2805 TGAATTTTGTCCCAAGGAATGGAGAGATTTAGCTATAGGATTCATTTGTTGATTGGAG 2864  
QY 2642 GAGCAACCACTTCAAAAAACCCACACAGAGTGTAAATAGCTCCGAGATACAGTGACCTG 2701  
Db 2865 GAGCAACCACTTCAAAAAACCCACACAGAGTGTAAATAGCTCCGAGATACAGTGACCTG 2924  
QY 2702 TAAATCCATGCTCTGAGCGCTCCAAAGTGTGGTGTGTTCCAGCTGTTTAGATGAAA 2761

```
Db 2925 TAATCAATGTCCTGGAGCGCGTCCAAGAGTGTGGTGTGTGTTCCAGCTGTTAGTAGAAA 2984
Qy 2762 ATCTAAGGATGAATATCTTTGAGGAAATCATGAAGAAATATGAAGATATTTAGACAGACC 2821
Db 2985 ATCTAAGGATGAATATCTTTGAGGAAATCATGAAGAAATATGAAGATATTTAGACAGACC 3044
Qy 2822 ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAACTCAAGCCAGAAAAAGTGTT 2881
Db 3045 ATTATGAGTCTCTCAAGGAGAGAGATACCTTAACTCAAGCCAGAAAAAGTGTT 3104
Qy 2882 TCCAAATGGATTGGCTGTCTGAACCTCACAGTGAAGCCACGTTTATTGGACCCAGG 2941
Db 3105 TCCAAATGGATTGGCTGTCTGAACCTCACAGTGAAGCCACGTTTATTGGACCCAGG 3164
Qy 2942 TCTTTGAAGCTATGACTGCAGAGCTGCAGAGCTGTGTGACTACATTTGACTGAAGCCTTTCTTTG 3001
Db 3165 TCTTTGAAGCTATGACTGCAGAGCTGTGTGACTACATTTGACTGAAGCCTTTCTTTG 3224
Qy 3002 ATGCTGGCAGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAGCACA 3061
Db 3225 ATGCTGGCAGCTCCGGGGCAAGTACCAGATCGAGGCTTCCCAAGATATTTAAGCACA 3284
Qy 3062 AAACAGTAGCTGGAGAGCCAGAGAGGTCTACGATGATGCCACAAATATCTGAACACAC 3121
Db 3285 AAACAGTAGCTGGAGAGCCAGAGAGGTCTACGATGATGCCACAAATATCTGAACACAC 3344
Qy 3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGTCTTGCCACAGCAGAGTA 3181
Db 3345 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTGGTTGGTCTTGCCACAGCAGAGTA 3404
Qy 3182 TCCAGAGCAGATTCACCTGTACGACAGAGCTGTGTCGCCAGGCTGCAGAGCCCATAG 3241
Db 3405 TCCAGAGCAGATTCACCTGTACGACAGAGCTGTGTCGCCAGGCTGCAGAGCCCATAG 3464
Qy 3242 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCATACT 3301
Db 3465 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCATACT 3524
Qy 3302 ACTGCCCTCAGACTTCATGCTCCCTTGATCTGGCATTCGGTACTACTGGGCCCTGT 3361
Db 3525 ACTGCCCTCAGACTTCATGCTCCCTTGATCTGGCATTCGGTACTACTGGGCCCTGT 3584
Qy 3362 TTGCGGTGCTGCTTTGGGTTAGAAGCTGACAGGCTATGAGGATGATGGTCACG 3421
Db 3585 TTGCGGTGCTGCTTTGGGTTAGAAGCTGACAGGCTATGAGGATGATGGTCACG 3644
Qy 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAGAAG 3481
Db 3645 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAGAAG 3704
Qy 3482 AGCTCCATGAAGATTGCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541
Db 3705 AGCTCCATGAAGATTGCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3764
Qy 3542 TCGCAGACCTGCGAAGTTGGGTACAGGGCATCCGCGGCTCTGGCTACCCACGACC 3601
Db 3765 TCGCAGACCTGCGAAGTTGGGTACAGGGCATCCGCGGCTCTGGCTACCCACGACC 3824
Qy 3602 AGCCCGACACACCGAGAGCTCACCAATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
Db 3825 AGCCCGACACACCGAGAGCTCACCAATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3884
Qy 3662 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGTCTTACAGAGTCTCAGGCCCTACT 3721
Db 3885 GCATTAGGTTAAAGAAATCATTAGCAATGGCACTGTCTTACAGAGTCTCAGGCCCTACT 3944
Qy 3722 TCTCCATTTGAAGTCCAAATATTTTGTCTGTGGGAGATTTCCAAAGATCAGGTTGAGG 3781
Db 3945 TCTCCATTTGAAGTCCAAATATTTTGTCTGTGGGAGATTTCCAAAGATCAGGTTGAGG 4004
Qy 3782 ATTATGCAATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAACCCATT 3841
Db 4005 ATTATGCAATTCAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAACCCATT 4064
```

```
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 3875
Db 4065 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 4098
```

## RESULT 3

AEA63658  
ID AEA63658 standard; DNA; 7122 BP.

XX AEA63658;

DT 08-SEP-2005 (first entry)

XX Methionine synthase, MTR, DNA sequence, SEQ ID 2.

XX Neuroleptic; Nootropic; Tranquillizer; folate metabolism;  
KW psychiatric disorder; schizophrenia; autism;  
KW attention deficit hyperactivity disorder; obsessive-compulsive disorder;  
KW gene; ds; Methionine synthase; chromosome 1.

XX Homo sapiens.

OS US6912492-B1.

PN 28-JUN-2005.

XX 23-MAY-2000; 2000US-00577266.

XX 25-MAY-1999; 99US-0136198P.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Johnson WG, Stenroos ES;

PI WPI; 2005-442724/45.

XX Estimating the susceptibility of an individual to have offspring that develop a developmental disorder comprises analyzing the nucleic acids and/or proteins for genes involved in folate, pyridoxine, and/or cobalamin metabolism.

PS Disclosure; SEQ ID NO 2; 84pp; English.

XX The present invention relates to a method (M1) for estimating the susceptibility of an individual to have offspring that develop a developmental disorder. The method comprises analyzing the nucleic acids and/or proteins from the biological sample, where analyzing results in a partial or full genotype for the alleles of two or more genes involved in folate, pyridoxine, and/or cobalamin metabolism. The present sequence is one such gene involved in folate metabolism. The method is also useful for diagnosing, preventing, and treating developmental disorders, e.g. schizophrenia, autism, attention deficit hyperactivity disorder, or obsessive-compulsive disorder. The protein encoded by the present sequence is also known as 5-methyltetrahydrofolate-homocysteine S-methyltransferase. The present sequence is located on chromosome 1q43.

XX Sequence 7122 BP; 2112 A; 1450 C; 1728 G; 1832 T; 0 U; 0 Other;

Query Match 89.7%; Score 3517; DB 14; Length 7122;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GTACCTGTGGAGAGCAGCTTCTCTGCGCGCCCTTGGGAGAGGAGACTCGACA 61

Db 225 GTACCTGTGGAGAGCAGCTTCTCTGCGCGCCCTTGGGAGAGGAGACTCGACA 284

Qy 62 ACATGTACCCGCGCTCCAGACCTGTGCAACCCGAGGTCTGAAGAAACCTCGGG 121

Db 285 ACATGTACCCGCGCTCCAGACCTGTGCAACCCGAGGTCTGAAGAAACCTCGGG 344

Qy 122 ATGAGATCAATGCCATTCTCTGCAAGAGGATTTATGCTGATGGAGGATGGGACCA 181

Db 345 ATGAGATCAATGCCATTCTGCAGAGAGGATTATGGTGTGGATGGGGATGGGGACCA 404  
Qy 182 TGAATCAGCGGGAAGAGCTAAACGAAGAACACCTTCCGAGGTGAGGAATTTAAAGATCATG 241  
Db 405 TGAATCAGCGGGAAGAGCTAAACGAAGAACACCTTCCGAGGTGAGGAATTTAAAGATCATG 464  
Qy 242 CCAGGCGGCTGAAGGCAACAAATGACATTTTAAGTATACTAGCTAGCGCTGATGCTATTACC 301  
Db 465 CCAGGCGGCTGAAGGCAACAAATGACATTTTAAGTATACTAGCTAGCGCTGATGCTATTACC 524  
Qy 302 AAATCCATAAGGAATATTCTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTTAGCA 361  
Db 525 AAATCCATAAGGAATATTCTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTTAGCA 584  
Qy 362 GCATAGTATTGCCAAGCTGACTATGGCTTGAAACATTTGGCTTACCGGATGAAACATGT 421  
Db 585 GCACTAGTATTGCCAAGCTGACTATGGCTTGAAACATTTGGCTTACCGGATGAAACATGT 644  
Qy 422 GCTCTCAGAGTGGCCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA 481  
Db 645 GCTCTCAGAGTGGCCAGAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAGA 704  
Qy 482 GGTTCGTGGCAGGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 705 GGTTCGTGGCAGGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCATCTGTGG 764  
Qy 542 AAAGCCGGANTATAGGAACATACATTTGATGAGCTTGTGAAGCATACCAAGACGAG 601  
Db 765 AAAGCCGGANTATAGGAACATACATTTGATGAGCTTGTGAAGCATACCAAGACGAG 824  
Qy 602 CCAAGGACTCTCGGATGGGGTGTGATCTTACTCATGGAACATAATTTTGATCTG 661  
Db 825 CCAAGGACTCTCGGATGGGGTGTGATCTTACTCATGGAACATAATTTTGATCTG 884  
Qy 662 CCAATGCCAAGGAGCGCTTGTTCGACTCCAAATCTTTTGAAGGAGAAATATGCTCCGC 721  
Db 885 CCAATGCCAAGGAGCGCTTGTTCGACTCCAAATCTTTTGAAGGAGAAATATGCTCCGC 944  
Qy 722 GGCCTATCTTTATTTAGGAGCGATCGTTGATAAAGAGTGGGCGGACTCTTTCCGACAGA 781  
Db 945 GGCCTATCTTTATTTAGGAGCGATCGTTGATAAAGAGTGGGCGGACTCTTTCCGACAGA 1004  
Qy 782 CAGGAGAGGNTTGTTCATCAGCGTCTCATGAGAACCACTCTCGATGATTAATTT 841  
Db 1005 CAGGAGAGGNTTGTTCATCAGCGTCTCATGAGAACCACTCTCATGAGAACCACTCTCATGAGNTTAAATTT 1064  
Qy 842 GTGCTTTGGTGCAGCTGAGATCAGACCTTTTATGAAATTAATTTGAAATGTACACAG 901  
Db 1065 GTGCTTTGGTGCAGCTGAAATGAGACCTTTTATGAAATTAATTTGAAATGTACACAG 1124  
Qy 902 CCTATGCTCTGTTATCCCAATGSCAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 961  
Db 1125 CCTATGCTCTGTTATCCCAATGSCAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 1184  
Qy 962 CGCCTTCTATGATGGCAGACCTAAAGATTTTGTCTATGATGAGTGGCTGGTCAATATAG 1021  
Db 1185 CGCCTTCTATGATGGCAGACCTAAAGATTTTGTCTATGATGAGTGGCTGGTCAATATAG 1244  
Qy 1022 TTGGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1245 TTGGAGATGCTGTGGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1304  
Qy 1082 ATTGTAAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGTCTGGTC 1141  
Db 1305 ATTGTAAGCCTTAGAGTTCCACCTGCCACTGCTTTTGAAGACATATGTTACTGTCTGGTC 1364  
Qy 1142 TAGAGCCCTTCAGGATTTGAGCCGTACACCAATTTTGTGTTAACTTGAAGAGCGCTGTAATG 1201  
Db 1365 TAGAGCCCTTCAGGATTTGAGCCGTACACCAATTTTGTGTTAACTTGAAGAGCGCTGTAATG 1424  
Qy 1202 TTGCGAGNTCAGGAATTTGCTTAACCTCATCTATGCGAGGAACCTATGAGAGCCCTGCT 1261  
Db 1425 TTGCGAGNTCAGGAATTTGCTTAACCTCATCTATGCGAGGAACCTATGAGAGCCCTGCT 1484

Qy 1262 GTGTTGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGGATGATG 1321  
Db 1485 GTGTTGCCAAAGTGCAGGTGGAAATGGGAGCCAGGTGTTGGATGTCAACATGGATGATG 1544  
Qy 1322 GCATGCTAGATGTCCTCAAGTGAATGACAGATTTTGAACCTTAATTTGCTTCCGAGCCAG 1381  
Db 1545 GCATGCTAGATGTCCTCAAGTGAATGACAGATTTTGAACCTTAATTTGCTTCCGAGCCAG 1604  
Qy 1382 ACATCGAAAGGTACCTTTTGTGCATCGACTCTCTCCATTTTGTGTGATTTGAAGCTGGGT 1441  
Db 1605 ACATCGAAAGGTACCTTTTGTGCATCGACTCTCTCCATTTTGTGTGATTTGAAGCTGGGT 1664  
Qy 1442 TAAAGTGTGCCAAGGAAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAGGAGAGAGCG 1501  
Db 1665 TAAAGTGTGCCAAGGAAAGTGCATTTGTCAATAGCATTTAGTCTGAAGGAGGAGAGAGCG 1724  
Qy 1502 ACTTCTTGAGAGAGCGCAGGAAGATTAAAGAGTATGGAGCTGCTATGGTGGTTCATGGCTT 1561  
Db 1725 ACTTCTTGAGAGAGCGCAGGAAGATTAAAGAGTATGGAGCTGCTATGGTGGTTCATGGCTT 1784  
Qy 1562 TTGATGAAGAGGACAGGCAACAGAACACAGACACAAAAATCAGAGTGTGCACCCGGGCT 1621  
Db 1785 TTGATGAAGAGGACAGGCAACAGAACACAGACACAAAAATCAGAGTGTGCACCCGGGCT 1844  
Qy 1622 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGACCCCTAATA 1681  
Db 1845 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGACCCCTAATA 1904  
Qy 1682 TCCTAACCATCTGGGACTGGAATGGGAGAACACAACTTGTATGCCATTTATTTTATCCATG 1741  
Db 1905 TCCTAACCATCTGGGACTGGAATGGGAGAACACAACTTGTATGCCATTTATTTTATCCATG 1964  
Qy 1742 CAAACAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 1801  
Db 1965 CAAACAAAGTCAATTAAGAAACATTAACCTGGAGCCAGAAATAAGTGGAGGTCTTTTCCAACT 2024  
Qy 1802 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTTCGAGAACCAATGCAATGGGTTTTTCCCTT 1861  
Db 2025 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTTTCGAGAACCAATGCAATGGGTTTTTCCCTT 2084  
Qy 1862 ACCATGCAATCAAGTCTGGCATGGAGTGTGAGTGTGAAACCTTCCCTGTGT 1921  
Db 2085 ACCATGCAATCAAGTCTGGCATGGAGTGTGAGTGTGAAACCTTCCCTGTGT 2144  
Qy 1922 ATGATGATATCCATTAAGGAACCTTTTCGAGCTCTGTGAAAGATCTCATCTGGAATAAGAGAC 1981  
Db 2145 ATGATGATATCCATTAAGGAACCTTTTCGAGCTCTGTGAAAGATCTCATCTGGAATAAGAGAC 2204  
Qy 1982 CTGAGGCCACTGAGAGCTTTTACGTTATGCCAGACTCAAGGCAACAGAGGGAAGAAAG 2041  
Db 2205 CTGAGGCCACTGAGAGCTTTTACGTTATGCCAGACTCAAGGCAACAGAGGGAAGAAAG 2264  
Qy 2042 TCATTCAGACTCATGATGGAGAAATGGCCCTGTGAGAGAACCTTGTGATGATGCCCTTG 2101  
Db 2265 TCATTCAGACTCATGATGGAGAAATGGCCCTGTGAGAGAACCTTGTGATGATGCCCTTG 2324  
Qy 2102 TGAAGGGCAATTCAAAAACATATTTTGAAGTACTGAGGAAGCCAGGTTTAAACCAAAAA 2161  
Db 2325 TGAAGGGCAATTCAAAAACATATTTTGAAGTACTGAGGAAGCCAGGTTTAAACCAAAAA 2384  
Qy 2162 AATATCCCCGACCTCTCAATATTAATTAAGAGCCCTGTATGATGGAATGGAATTTGTTG 2221  
Db 2385 AATATCCCCGACCTCTCAATATTAATTAAGAGCCCTGTATGATGGAATGGAATTTGTTG 2444  
Qy 2222 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2281  
Db 2445 GTGATCTTTTGGAGCTGGAATAATTTTCTACCTCAGGTTTAAAGTCAAGCCCGGTTA 2504  
Qy 2282 TGAAGAGGCTCTGTGGCCACCTTTATCCCTTTTCATATGGAAGAAAGAGAGAAACAGAG 2341  
Db 2505 TGAAGAGGCTCTGTGGCCACCTTTATCCCTTTTCATATGGAAGAAAGAGAGAAACAGAG 2564

QY 2342 TGCTTAACGGCACAGTAGAAGAGAGACCCCTTACAGGGCCACATCGTGTGCGCACTG 2401  
DB |||||  
QY 2565 TGCTTAACGGCACAGTAGAAGAGAGAGACCCCTTACAGGGCCACATCGTGTGCGCACTG 2624  
DB |||||  
QY 2402 TTAAGGCGCAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCCCTTGGCTGCAATA 2461  
DB |||||  
QY 2625 TTAAGGCGCAGCTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCCCTTGGCTGCAATA 2684  
DB |||||  
QY 2462 ATTTCGAGTTATTGATTTTAGAGTATGATCTCCATGTGTATAGATACTGAAAGCTGCTC 2521  
DB |||||  
QY 2685 ATTTCGAGTTATTGATTTTAGAGTATGATCTCCATGTGTATAGATACTGAAAGCTGCTC 2744  
DB |||||  
QY 2522 TTGACCAACAGCAGATATATTTGGCTGTTCAGAGCTCATCACTCCCTCCCTCGATGAA 2581  
DB |||||  
QY 2745 TTGACCAACAGCAGATATATTTGGCTGTTCAGAGCTCATCACTCCCTTCCTGGATGAA 2804  
DB |||||  
QY 2582 TGATTTTGTGTGCAAGGAATGGAGAGATTAGCTATAAGGATTCCAATTGCTGATGGAG 2641  
DB |||||  
QY 2805 TGATTTTGTGTGCAAGGAATGGAGAGATTAGCTATAAGGATTCCAATTGCTGATGGAG 2864  
DB |||||  
QY 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701  
DB |||||  
QY 2865 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG 2924  
DB |||||  
QY 2702 TAATCCATGTCTGGAGCGCTCCAAGAGTGTGGTGTCTCCAGCTGTAGATGAAA 2761  
DB |||||  
QY 2925 TAATCCATGTCTGGAGCGCTCCAAGAGTGTGGTGTCTCCAGCTGTAGATGAAA 2984  
DB |||||  
QY 2762 ATCTAAGGATGAATCTTTGAGGAATCATGGAAGAAATAGAGATATTAGACAGGACC 2821  
DB |||||  
QY 2985 ATCTAAGGATGAATCTTTGAGGAAATCATGGAAGAAATAGAGATATTAGACAGGACC 3044  
DB |||||  
QY 2822 ATTATGAGTCTCAAGGAGAGAGATACCTTAAGCTCAAGCCAGAAAAGTGTT 2881  
DB |||||  
QY 3045 ATTATGAGTCTCAAGGAGAGAGATACCTTAAGCTCAAGCCAGAAAAGTGTT 3104  
DB |||||  
QY 2882 TCCAAATGGATTTGGCTGTGAACTCACCAGTGAAGCCACGTTTATTGGGACCCAGG 2941  
DB |||||  
QY 3105 TCCAAATGGATTTGGCTGTGAACTCACCAGTGAAGCCACGTTTATTGGGACCCAGG 3164  
DB |||||  
QY 2942 TCTTTGAAGACTATGACTGCAGAACTGGTGGACTACATTTGACTGGAAGCCTTTCTTTG 3001  
DB |||||  
QY 3165 TCTTTGAAGACTATGACTGCAGAACTGGTGGACTACATTTGACTGGAAGCCTTTCTTTG 3224  
DB |||||  
QY 3002 ATGCTGGCAGCTCCGGGCAAGTACCCGAATCGAGAGCTTCCCAAGATATTTAACGACA 3061  
DB |||||  
QY 3225 ATGCTGGCAGCTCCGGGCAAGTACCCGAATCGAGAGCTTCCCAAGATATTTAACGACA 3284  
DB |||||  
QY 3062 AAACAGTAGGTGGAGAGCCAGGAAGCTCTACGATGATGCCCAATATGCTGAACACAC 3121  
DB |||||  
QY 3285 AAACAGTAGGTGGAGAGCCAGGAAGCTCTACGATGATGCCCAATATGCTGAACACAC 3344  
DB |||||  
QY 3122 TGATTAGTCAAAAAGAACTCCGGGCCCGGGGTGGTGGTCTTGCGCCAGCAAGAGTA 3181  
DB |||||  
QY 3345 TGATTAGTCAAAAAGAACTCCGGGCCCGGGGTGGTGGTCTTGCGCCAGCAAGAGTA 3404  
DB |||||  
QY 3182 TCCAAGACGACATTCACCTGTAACGAGAGCTGTGTGCCCCAGGCTGCGAGAGCCCATAG 3241  
DB |||||  
QY 3405 TCCAAGACGACATTCACCTGTAACGAGAGCTGTGTGCCCCAGGCTGCGAGAGCCCATAG 3464  
DB |||||  
QY 3242 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCATACT 3301  
DB |||||  
QY 3465 CCACCTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGACCGAGGCATACT 3524  
DB |||||  
QY 3302 ACTGCTCTCAGACTTTCATCCCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCCCTGT 3361  
DB |||||  
QY 3525 ACTGCTCTCAGACTTTCATCCCTCCCTTGCATTTCTGGCATCCGTGACTACCTGGGCCCTGT 3584  
DB |||||  
QY 3362 TTGGCGTTCCTGCTTTGGGTTAGAGAGCTGAGAGGCCCTATGAGATGATGGTGAACG 3421  
DB |||||  
QY 3585 TTGGCGTTCCTGCTTTGGGTTAGAGAGCTGAGCAAGGCCCTATGAGATGATGGTGAACG 3644  
DB |||||  
QY 3422 ACTACGACGACATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAAG 3481

DB |||||  
QY 3645 ACTACGACGACATCATGGTCAAGGGCGCTGGGGACCGGCTGGCAGAGGCTTTGCAGAAC 3704  
DB |||||  
QY 3482 AGCTCATGAAGAGTTCCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCTGGACG 3541  
DB |||||  
QY 3705 AGCTCATGAAGAGTTCCGCCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCTGGACG 3764  
DB |||||  
QY 3542 TCGCAGACCTGCGAAGTTGCGGTACAAGGGCATCCGCCCGGCTCCTGGCTACCCAGCC 3601  
DB |||||  
QY 3765 TCGCAGACCTGCGCAGGCTGCGGTACAAGGGCATCCGCCCGGCTCCTGGCTACCCAGCC 3824  
DB |||||  
QY 3602 AGCCCCACACACCCGAGAAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
DB |||||  
QY 3825 AGCCCCACACACCCGAGAAAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3884  
DB |||||  
QY 3662 GCATTAGGTTAAACAGAAATCATTAGCAATGGCAGCTCCTCTCAGCAGTCTCAGGCTCTACT 3721  
DB |||||  
QY 3885 GCATTAGGTTAAACAGAAATCATTAGCAATGGCAGCTCCTCTCAGCAGTCTCAGGCTCTACT 3944  
DB |||||  
QY 3722 TCTCCCAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTTCCAAGGATCAGGTTGAGG 3781  
DB |||||  
QY 3945 TCTCCCAATTTGAAGTCCAAATATTTTGTGCTGGGGAAGATTTTCCAAGGATCAGGTTGAGG 4004  
DB |||||  
QY 3782 ATTATGCATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 3841  
DB |||||  
QY 4005 ATTATGCATTTGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCATTT 4064  
DB |||||  
QY 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 3875  
DB |||||  
QY 4065 TGGGATATGATACAGACTAACTTTTTTTTTTTTTTTT 4098  
DB |||||

## RESULT 4

AAF31101  
ID AAF31101 standard; cDNA; 7224 BP.

XX AC AAF31101;

XX DT 27-APR-2001 (first entry)

XX DE Methionine synthase coding sequence #2.

XX KW Analyte-binding enzyme; analyte analysis; ss.

XX OS Homo sapiens.

XX PN WO200102600-A2.

XX PD 11-JAN-2001.

XX PF 30-JUN-2000; 2000WO-US018057.

XX PR 06-JUL-1999; 99US-00347878.

XX PR 06-DEC-1999; 99US-00457205.

XX PA (GEAT ) GEN ATOMICS.

XX PI Yuan C;

XX DR WPI; 2001-071583/08.

XX PT Assaying method, useful for prognosis and diagnosis of disease, comprises contacting sample with a mutant analyte-binding enzyme and detecting binding.

XX PS Disclosure; Page; 187pp; English.

XX CC The present invention relates to a method for assaying an analyte in a sample comprising: contacting the sample with a mutant analyte-binding enzyme which has binding affinity for the analyte or an immediate analyte enzymatic conversion product but has attenuated catalytic activity; and detecting resulting binding. The method is useful in monitoring biological systems/processes, or prognosis/diagnosis of disease caused by





QY	2042	TCATTTCAGATGTGATGGAGAAATGGCCCTGTCGAGAAACGCTTGAGTATGCCCTTG	2101
Db	2373	TCATTTCAGATGTGATGGAGAAATGGCCCTGTCGAGAAACGCTTGAGTATGCCCTTG	2432
QY	2102	TGAAGGCAATTGAAAAACATATTATTGAGGATACCTGAGGAAGCCAGGTAAACCAAAAA	2161
Db	2433	TGAAGGCAATTGAAAAACATATTATTGAGGATACCTGAGGAAGCCAGGTAAACCAAAAA	2492
QY	2162	AATATCCCGACCTCTCAATATATATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2221
Db	2493	AATATCCCGACCTCTCAATATATATTGAAGGACCCCTGATGAATGGAATGAAATTTGTTG	2552
QY	2222	GTGATCTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTTAAAGTTCAGCCCGGTTA	2281
Db	2553	GTGATCTTTTGGAGCTGGAATAATGTTTCTACTCAGGTTTAAAGTTCAGCCCGGTTA	2612
QY	2282	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAAACACAGAG	2341
Db	2613	TGAAGAGGCTGTTGGCCACCTTATCCCTTTTCATGGAAGAAAGAGAAACACAGAG	2672
QY	2342	TGCTTAACGGCACAGTATAGAAAGAGAACCCCTTACCAAGGACCATCGTCTGGCCACTG	2401
Db	2673	TGCTTAACGGCACAGTATAGAAAGAGAACCCCTTACCAAGGACCATCGTCTGGCCACTG	2732
QY	2402	TTAAAGGGAGCTGCACGACATAGGCAAGACATAGTTGGAGTAGTCTTTGGCTGCAATA	2461
Db	2733	TTAAAGGGAGCTGCACGACATAGGCAAGACATAGTTGGAGTAGTCTTTGGCTGCAATA	2792
QY	2462	ATTTCCAGTTATTGATTTAGGAGTCATGACTCCATGTAAGATACGAAAGCTGCTC	2521
Db	2793	ATTTCCAGTTATTGATTTAGGAGTCATGACTCCATGTAAGATACGAAAGCTGCTC	2852
QY	2522	TTGACCAAAAGCAGATATAATTGGCCCTGTGAGGACTCATCCTCTCCCTGGATGAAA	2581
Db	2853	TTGACCAAAAGCAGATATAATTGGCCCTGTGAGGACTCATCCTCTCCCTGGATGAAA	2912
QY	2582	TGATTTTGTGTCGAAGAAATGAGAGATTAGCTATAAGGATTCATTTGTTGATGGAG	2641
Db	2913	TGATTTTGTGTCGAAGAAATGAGAGATTAGCTATAAGGATTCATTTGTTGATGGAG	2972
QY	2642	GAGCAACACCTTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	2701
Db	2973	GAGCAACACCTTTCAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGCACCTG	3032
QY	2702	TAATCCATGTCTCGACGCGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA	2761
Db	3032	TAATCCATGTCTCGACGCGTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA	3092
QY	2762	ATCTAAAGGATGAATACCTTGGAGGAATCATGGAAGATATGAGATATTAGACGAGCC	2821
Db	3093	ATCTAAAGGATGAATACCTTGGAGGAATCATGGAAGATATGAGATATTAGACGAGCC	3152
QY	2822	ATTATGAGTCTCTCAAGGAGGAGATACCTTACCTTAAAGTCAAGCCAGAAAGTGGTT	2881
Db	3153	ATTATGAGTCTCTCAAGGAGGAGATACCTTACCTTAAAGTCAAGCCAGAAAGTGGTT	3212
QY	2882	TCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACCGTTTATTGGGACCCAGG	2941
Db	3213	TCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACCGTTTATTGGGACCCAGG	3272
QY	2942	TCCTTGAAGACTATGACCTCGAGAGCTGGTGGACTACATTTGACTGGAGCCTTCTTTG	3001
Db	3273	TCCTTGAAGACTATGACCTCGAGAGCTGGTGGACTACATTTGACTGGAGCCTTCTTTG	3332
QY	3002	ATGCTGTCAGCTCCGGGCGAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA	3061
Db	3333	ATGCTGTCAGCTCCGGGCGAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAACGACA	3392
QY	3062	AAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCACAAATATGCTGAACACAC	3121
Db	3393	AAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCACAAATATGCTGAACACAC	3452

QY	3122	TGATTAGTCAAAAGAAACTCCGGGCCCGGGGTGTGGTTGGGTCTTGCCGACACAGAGTA	3181
Db	3453	TGATTAGTCAAAAGAAACTCCGGGCCCGGGGTGTGGTTGGGTCTTGCCGACACAGAGTA	3512
QY	3182	TCCAAGACCAATTCACCTGTACGAGAGGCTGTGTGTGCCCAAGGCTGCAGAGCCCATAG	3241
Db	3513	TCCAAGACCAATTCACCTGTACGAGAGGCTGTGTGTGCCCAAGGCTGCAGAGCCCATAG	3572
QY	3242	CCACTTCTCTATGGGTTAAGGCAACAGGCTGAGAAAGACTCTGCCAGACGGAGCCATCT	3301
Db	3573	CCACTTCTCTATGGGTTAAGGCAACAGGCTGAGAAAGACTCTGCCAGACGGAGCCATCT	3632
QY	3302	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCCTGCCATCCGCTGACTACCTGGGCTGT	3361
Db	3633	ACTGCCCTCTCAGACTTCATCGCTCCCTTGCATTCCTGCCATCCGCTGACTACCTGGGCTGT	3692
QY	3362	TTGCCCTTCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG	3421
Db	3693	TTGCCCTTCCCTGTCTTTGGGGTAGAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG	3752
QY	3422	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAGAG	3481
Db	3753	ACTACAGCAGCATCATGGTCAAGGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAGAG	3812
QY	3482	AGCTCCATGAAGAGTTCGCGCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGACG	3541
Db	3813	AGCTCCATGAAGAGTTCGCGCGAGAACTGTGGGCTTACTGTGGCAGTGAAGCAGCTGACG	3872
QY	3542	TGCGAGACCTGCGAAGGTTTCGGTCAAGGGCATCCGCGGCTCTCTGGCTACCCGAGCC	3601
Db	3873	TGCGAGACCTGCGCGAGGCTTCGGTCAAGGGCATCCGCGGCTCTCTGGCTACCCGAGCC	3932
QY	3602	AGCCGACACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3661
Db	3933	AGCCGACACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG	3992
QY	3662	GCATTAGGTTTAAACAGAACTCATTAGCAATGCGACCTGCTTCAGCAGTCTCAGGCTCTACT	3721
Db	3993	GCATTAGGTTTAAACAGAACTCATTAGCAATGCGACCTGCTTCAGCAGTCTCAGGCTCTACT	4052
QY	3722	TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTTCAAAGGATCAGGTTGAGG	3781
Db	4053	TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTTCAAAGGATCAGGTTGAGG	4112
QY	3782	ATTATGCAATGAGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAGCCCATTT	3841
Db	4113	ATTATGCAATGAGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGAGCCCATTT	4172
QY	3842	TGGGATATGATACAGACTAACTTT	3875
Db	4173	TGGGATATGATACAGACTAACTTT	4206

RESULT 5  
ADI39109  
ID ADI39109 standard; DNA; 3798 BP.  
XX  
AC ADI39109;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Homo sapiens meth DNA.  
XX  
KW fermentation; methionine; Coryneform bacterium; methionine synthase;  
KW Meth; animal feed additive; sulphur; human nutrition; animal nutrition;  
KW cosmetic; pharmaceutical; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS 1. 3798  
FT / \*tag= a  
FT / product= "meth"







Db 3601 ATTAGGTTAACAGATCAITAGCAATGGGACCTGCTTCAGAGCTCAGGCGCTCTACTTC 3660

Qy 3724 TCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAGATTTCCAAAGATCAGGTTGAGGAT 3783

Db 3661 TCCAAATTTGAAGTCCAAATATTTTGTGTGGGAAGATTTCCAAAGATCAGGTTGAGGAT 3720

Qy 3784 TATGCAATTTGAGGAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCCATTTCG 3843

Db 3721 TATGCAATTTGAGGAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGACCCATTTCG 3780

Qy 3844 GGATATGATACAGACTAA 3861

Db 3781 GGATATGATACAGACTAA 3798

RESULT 6

AAV34063

ID AAV34063 standard; cDNA; 3920 BP.

XX AAV34063;

DT 01-MAR-1999 (first entry)

XX Human methionine synthase cDNA (wild-type).

XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect; cardiovascular disease; colon cancer; diagnosis; therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 64..3861

FT /\*tag= a

FT /EC\_number= "2.1.1.13"

XX CA2217153-A.

XX 27-MAY-1998.

XX 27-NOV-1997; 97CA-02217153.

XX 27-NOV-1996; 96US-0031964P.

XX 20-JUN-1997; 97US-0050310P.

XX (MART-) MARTINEX R & D INC.

XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;

XX WPI; 1998-569373/49.

XX P-PSDB; AAW68594.

XX DNA encoding methionine synthase polypeptide - and corresponding

XX polypeptide, cells, antibody and therapeutic methods.

XX Claim 6; Fig 3; 64pp; English.

XX This cDNA clone encodes human wild-type methionine synthase (see

XX AAW68594). A survey of the NCBI databases yielded several sequences

XX corresponding to methionine synthases from different organisms.

XX Comparison of these sequences generated 4 very conserved regions.

XX Degenerate oligonucleotides (see AAV34064-104) based on these conserved

XX sequences were used as primers for RT-PCR with human and mouse mRNA. PCR

XX products were subcloned, sequenced and aligned. Additional sequences were

XX subsequently obtained by further PCR and inverse PCR. The full-length

XX sequence encodes a 1265-amino acid protein of predicted mol.wt. 141 kDa.

XX Using FISH, the gene was mapped to human chromosome band 1q43. Mutations

XX in this gene are associated with hyperhomocysteinaemia. One missense

XX mutation (2758 C to G) and a 3 bp deletion (2640-42) have been identified

XX in patients of the cblG complementation group. A polymorphism (2756 A to

XX G) has also been detected. The invention relates to the diagnosis and

XX treatment of patients at risk for methionine synthase deficiency, and

XX associated altered risk for neural tube defects, cardiovascular disease,

XX and colon cancer. Methods for detecting sequence variants for methionine

CC synthase, involving SSCP analysis or RFLP polymorphism analysis, are

CC claimed, as are methods for screening for therapeutic compounds that

CC modulate methionine synthase activity

XX

Sequence 3920 BP; 1114 A; 866 C; 975 G; 965 T; 0 U; 0 Other;

Query Match 34.4%; Score 1349; DB 2; Length 3920;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 2649; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 GGTCACTCTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60

Db 1 GGTCACTCTGGAGAGCAGCTCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60

Qy 61 AACATGTCAACCCGCTCCAAGACCTGTGCGAACCCGAGGCTGAGAAAAACCTTCGCG 120

Db 61 AACATGTCAACCCGCTCCAAGACCTGTGCGAACCCGAGGCTGAGAAAAACCTTCGCG 120

Qy 121 GATGATCAATGCCATTTCTGCGAAGAGGATTTGGTCTGGATGGAGGATGGGACC 180

Db 121 GATGATCAATGCCATTTCTCCAGAAAGAGGATTTATGGTCTGGATGGAGGATGGGACC 180

Qy 181 ATGATCCAGCGGAGAGGCTTAAACGAGAACTTCCGAGGTCAGGAATTTAAAGATCAT 240

Db 181 ATGATCCAGCGGAGAGGCTTAAACGAGAACTTCCGAGGTCAGGAATTTAAAGATCAT 240

Qy 241 GCCAGCCGCTGAAAGGCAACAATGACATTTTAAAGTATAACTCAGCTCATGTCATTAC 300

Db 241 GCCAGCCGCTGAAAGGCAACAATGACATTTTAAAGTATAACTCAGCTCATGTCATTAC 300

Qy 301 CAAATCCATAAGGAATACCTTGTGCTGGGCGGAGATATCATTTGAAACAATATCTTTAGC 360

Db 301 CAAATCCATAAGGAATACCTTGTGCTGGGCGGAGATATCATTTGAAACAATATCTTTAGC 360

Qy 361 AGCCTAGTATTGCCCAAGCTGACTATGCTTTGAACACTTGGCCCTACCGATGAACATG 420

Db 361 AGCCTAGTATTGCCCAAGCTGACTATGCTTTGAACACTTGGCCCTACCGATGAACATG 420

Qy 421 TGCTCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAG 480

Db 421 TGCTCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAATTAAG 480

Qy 481 AGTTTGTGGCAGGGGCTCTGGGTCGACTAATAGACACTCTCTGTGTCCCATCTGTG 540

Db 481 AGTTTGTGGCAGGGGCTCTGGGTCGACTAATAGACACTCTCTGTGTCCCATCTGTG 540

Qy 541 GAAAGCCCGATTTATAGGAACATCAATTTGATGAGCTTGTGGAAGCAATACCAAGAGCAG 600

Db 541 GAAAGCCCGATTTATAGGAACATCAATTTGATGAGCTTGTGGAAGCAATACCAAGAGCAG 600

Qy 601 GCCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTACT 660

Db 601 GCCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTACT 660

Qy 661 GCCAATGCCAAGCGAGCTTTTGCATCTCCAAAATCTTTTGGAGGAGAAATATGTCCTC 720

Db 661 GCCAATGCCAAGCGAGCTTTTGCATCTCCAAAATCTTTTGGAGGAGAAATATGTCCTC 720

Qy 721 CGGCTCTCTTTTTCAGGAGGATCGTTGATTAAGTGGGCGGAGCTCTTCCGACAG 780

Db 721 CGGCTCTCTTTTTCAGGAGGATCGTTGATTAAGTGGGCGGAGCTCTTCCGACAG 780

Qy 781 ACAGGAGAGGATTTGTCTCATCGCGTGTCTCATGAGAACCACTCTGCAATGGATTAAT 840

Db 781 ACAGGAGAGGATTTGTCTCATCGCGTGTCTCATGAGAACCACTCTGCAATGGATTAAT 840

Qy 841 TGTGCTTTGGGTGACGCTGAGATGAGACTTTTATTTGAAATATTTGGAATAATGTACAACA 900

Db 841 TGTGCTTTGGGTGACGCTGAGATGAGACTTTTATTTGAAATATTTGGAATAATGTACAACA 900

Qy 901 GCTATGCTCTCTGTTATCCCAATGAGCTCTTCCCAACACCTTTGGTGCATATGATGAA 960

Db 901 GCTATGCTCTCTGTTATCCCAATGAGCTCTTCCCAACACCTTTGGTGCATATGATGAA 960

QY 961 ACGCTTCTATGATGGCCAGCACCTAAAGATTTTCTATGATGGCTTGGTCAATATA 1020  
Db 961 ACCCTTTCTATGATGGCCAGCACCTAAAGATTTTCTATGATGGCTTGGTCAATATA 1020  
QY 1021 GTTGGAGGATGCTGTGGGTCAAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
Db 1021 GTTGGAGGATGCTGTGGGTCAAACACAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
QY 1081 AATTGTAAGCTAGATTTCAACCTGCCA CTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
Db 1081 AATTGTAAGCTAGATTTCAACCTGCCA CTGCTTTTGAAGGACATATGTTACTGTCTGGT 1140  
QY 1141 CTAGAGCCCTCAGGATTTGACCGGTACACCACTTTTGAATTTGAGAGCGCTGTAAT 1200  
Db 1141 CTAGAGCCCTCAGGATTTGACCGGTACACCACTTTTGAATTTGAGAGCGCTGTAAT 1200  
QY 1201 GTTGCAGGATCAAGAAATTTGCTTAAACTCATATGCGAGAACTATGAAGAGCCTTG 1260  
Db 1201 GTTGCAGGATCAAGAAATTTGCTTAAACTCATATGCGAGAACTATGAAGAGCCTTG 1260  
QY 1261 TGTGTTGCCAAAGTGCAGGTGGAATTTGGAGCCCGAGGTGTTGGATGTCAACATGGATGAT 1320  
Db 1261 TGTGTTGCCAAAGTGCAGGTGGAATTTGGAGCCCGAGGTGTTGGATGTCAACATGGATGAT 1320  
QY 1321 GGCATGCTAGATGGTCCAAGTGCAGTACCAAGATTTTGCACCTTAATTTGCTTCCGAGCCA 1380  
Db 1321 GGCATGCTAGATGGTCCAAGTGCAGTACCAAGATTTTGCACCTTAATTTGCTTCCGAGCCA 1380  
QY 1381 GACATCCGAAAGTACCTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGG 1440  
Db 1381 GACATCCGAAAGTACCTTTGTGATCGACTCTCCCAATTTTGTGCTGATTTGAAGCTGGG 1440  
QY 1441 TTAAGTGTGCCAAGGGAAGTGCATTTGCTCAATAGCATTTAGTCTGAAGGAGGAGGAC 1500  
Db 1441 TTAAGTGTGCCAAGGGAAGTGCATTTGCTCAATAGCATTTAGTCTGAAGGAGGAGGAC 1500  
QY 1501 GACTTCTTGGAGAGCCGAGGAATTTAAAGATTTGAGCTGCTATGGTGGTCAATGGCT 1560  
Db 1501 GACTTCTTGGAGAGCCGAGGAATTTAAAGATTTGAGCTGCTATGGTGGTCAATGGCT 1560  
QY 1561 TTTGATGAAGAGGACAGGCAACAGAAACAGACACAAATTCAGAGTGTGCACCCGGGCC 1620  
Db 1561 TTTGATGAAGAGGACAGGCAACAGAAACAGACACAAATTCAGAGTGTGCACCCGGGCC 1620  
QY 1621 TACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGCACCTTAAT 1680  
Db 1621 TACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTTGCACCTTAAT 1680  
QY 1681 ATCTTAACCATTTGGAGCTGGATGGAGGACACAACTTGTATGCCATTTATTTATCCAT 1740  
Db 1681 ATCTTAACCATTTGGAGCTGGATGGAGGACACAACTTGTATGCCATTTATTTATCCAT 1740  
QY 1741 GCAACAAAGTCATTAAAGAAACATTACCTGGAGCAGAAATAGTGGAGGTCTTTTCCAAAC 1800  
Db 1741 GCAACAAAGTCATTAAAGAAACATTACCTGGAGCAGAAATAGTGGAGGTCTTTTCCAAAC 1800  
QY 1801 TTGCTCTTCTCTTCCGAGGAATGGAAGCCATTGAGAGCAATGCAATGGGGTTTCTCTT 1860  
Db 1801 TTGCTCTTCTCTTCCGAGGAATGGAAGCCATTGAGAGCAATGCAATGGGGTTTCTCTT 1860  
QY 1861 TACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTG 1920  
Db 1861 TACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTG 1920  
QY 1921 TATGATGATATCCATAAGGAATCTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 1980  
Db 1921 TATGATGATATCCATAAGGAATCTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAC 1980  
QY 1981 CCTGAGGCCACTGAGAAGCTCTTACGTTATGCCCCAGACTCAAGGCACAGGAGGAGGAAA 2040  
Db 1981 CCTGAGGCCACTGAGAAGCTCTTACGTTATGCCCCAGACTCAAGGCACAGGAGGAGGAAA 2040

QY 2041 GTCAATTGAGATGATGAGTGGAGAAATGGCCCTGTGCAAGAAACGCCTTGATGATGCCCCTT 2100  
Db 2041 GTCAATTGAGATGATGAGTGGAGAAATGGCCCTGTGCAAGAAACGCCTTGATGATGCCCCTT 2100  
QY 2101 GTGAAGGGCAATTGAAAAACATATTTATTTAGGATATCTGAGAAAGCCAGGTTAAACCAAAAA 2160  
Db 2101 GTCAAGGGCAATTGAAAAACATATTTATTTAGGATATCTGAGAAAGCCAGGTTAAACCAAAAA 2160  
QY 2161 AATATCCCGACCTCTCAATATATTTTGAAGGACCCCTGATGAATGGAATGAAAATTTGTT 2220  
Db 2161 AATATCCCGACCTCTCAATATATTTTGAAGGACCCCTGATGAATGGAATGAAAATTTGTT 2220  
QY 2221 GGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAGCCCGGGTT 2280  
Db 2221 GGTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAGTCAGCCCGGGTT 2280  
QY 2281 ATGAAGAAGGCTGTTGGCCACCTTATCCCTTTTCATGGAATAAGGAGAAAGAAACGAGA 2340  
Db 2281 ATGAAGAAGGCTGTTGGCCACCTTATCCCTTTTCATGGAATAAGGAGAAAGAAACGAGA 2340  
QY 2341 GTGCTTAAACGGCACAGTAGAAGGAGGACCCCTTACAGGGCACCATCGTGTGGCCACT 2400  
Db 2341 GTGCTTAAACGGCACAGTAGAAGGAGGACCCCTTACAGGGCACCATCGTGTGGCCACT 2400  
QY 2401 GTTAAAGGGCAGCTGCACGACATAGCAAGAAACATAGTTGGAGTAGTCTCTTGGCTGCAAT 2460  
Db 2401 GTTAAAGGGCAGCTGCACGACATAGCAAGAAACATAGTTGGAGTAGTCTCTTGGCTGCAAT 2460  
QY 2461 AATTTCCGAGTTATGATTTAGGAGTCATGATCTCCATGTGATGAATATCTGAAAAGCTGCT 2520  
Db 2461 AATTTCCGAGTTATGATTTAGGAGTCATGATCTCCATGTGATGAATATCTGAAAAGCTGCT 2520  
QY 2521 CTTGACACAAAGCAGATATATTTGGCTGCTGAGGACTCATCTCCTTCCCTGGATGAA 2580  
Db 2521 CTTGACACAAAGCAGATATATTTGGCTGCTGAGGACTCATCTCCTTCCCTGGATGAA 2580  
QY 2581 ATCAATTTTGTGTCGCAAGGAATGGAGATAGCTATAAGGATTCATTTGTTGATTGGA 2640  
Db 2581 ATCAATTTTGTGTCGCAAGGAATGGAGATAGCTATAAGGATTCATTTGTTGATTGGA 2640  
QY 2641 GGAGCAACCACTTCAAAAAACCCACACAGCAGTTAA 2675  
Db 2641 GGAGCAACCACTTCAAAAAACCCACACAGCAGTTAA 2675  
RESULT 7  
ADRO7169  
ID ADRO7169 standard; cdna; 4172 BP.  
XX  
AC ADRO7169;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Full length human cDNA useful for treating neurological disease Seq 675.  
XX  
KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;  
KW osteoporosis; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; dementia; short memory; cancer;  
KW sense or motor function; emotional reaction; fear response; panic;  
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
KW tranquiliser.  
XX  
OS Homo sapiens.  
XX  
FN EPI47413-A2.  
XX  
PD 18-AUG-2004.  
XX  
PF 12-FEB-2004; 2004BP-00003145.  
XX  
PR 14-FEB-2003; 2003JP-00102207.  
XX  
PR 09-MAY-2003; 2003JP-00131452.

```
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI: 2004-583265/57.
DR PSDB; ADR09125.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 675; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
XX molecules and the encoded proteins thereof. Specifically, it refers to
XX cDNA clones obtained by an oligo-capping method, where none of these
XX clones are identical to any known human mRNAs. The present invention
XX describes an immunosay to identify agonists and antagonists, as well as
XX antibodies, antisense molecules and siRNAs that can all be used to bind
XX to and modulate expression of the cDNA molecules. As such, these
XX molecules are useful for diagnostic markers or therapeutic targets for
XX the various diseases or morbid states. In particular, they are useful in
XX gene therapy for treating osteoporosis, neurological disease, Alzheimer's
XX disease, Parkinson's disease, dementia, short memory and various cancers,
XX as well as for maintaining equilibrium of sense or motor function, and
XX for treating emotional reaction, fear response and panic. Accordingly,
XX they exhibit osteopathic, neuroprotective, neurotropic, antiparkinsonian,
XX cytotatic and tranquiliser activities. This polynucleotide is a full
XX length human cDNA sequence of the invention. NOTE: This sequence is not
XX given in the sequence listing of the specification but can be obtained on
XX CD-ROM from the European Patent Office, Vienna Sub-office.
XX
XX Sequence 4172 BP; 1175 A; 835 C; 980 G; 1182 T; 0 U; 0 Other;
XX
XX Query Match 7.9%; Score 308; DB 13; Length 4172;
XX Best Local Similarity 99.5%; Pred. No. 9.2e-113;
XX Matches 408; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 3467 AGCCCTTTGCAGAGAGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTACTGTGGCA 3526
XX 1015 AGCCCTTTGCAGAGAGCTCCATGAAGAGTTCGCCGAGAACTGTGGGCTACTGTGGCA 1074
XX 3527 GTGAGCAGCTGGAGCTGGAGACCTGGGAAGTGTGGGTACAGGGGATCCGCCCGGCTC 3586
XX 1075 GTGAGCAGCTGGAGCTGGAGACCTGGGAAGTGTGGGTACAGGGGATCCGCCCGGCTC 1134
XX 3587 CTGCTACCCCGCCAGCCGACACACCGAGAGGCTCACCATGTGGAGACTCGCAGACA 3646
XX 1135 CTGCTACCCCGCCAGCCGACACACCGAGAGGCTCACCATGTGGAGACTCGCAGACA 1194
XX 3647 TCAGCAGCTCTACAGGCATTAGGTAAACAGAACTATTAGCAATGGCACCTGCTTCAGCAG 3706
XX 1195 TCAGCAGCTCTACAGGCATTAGGTAAACAGAACTATTAGCAATGGCACCTGCTTCAGCAG 1254
XX 3707 TCTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTCTGGGGAGATTTCGA 3766
XX 1255 TCTCAGGCTCTACTTCTCAATTTGAAGTCCAAATATTTTGTCTGGGGAGATTTCGA 1314
XX 3767 AGGATCAGGTTGAGGATTATGCAATTCAGGAAGACATATCTGCTGGGTTGAGAAAT 3826
XX 1315 AGGATCAGGTTGAGGATTATGCAATTCAGGAAGACATATCTGCTGGGTTGAGAAAT 1374
XX 3827 GGCTTGGACCCATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTT 3876
XX 1375 GGCTTGGACCCATTTTGGGATATGATACAGACTAACTTTTTTTTTTTTTTTTTT 1424
XX
XX RESULT 8
XX AAV88164
XX ID AAV88164 standard; cDNA; 305 BP.
XX
XX AC AAV88164;
XX
```

```
DT 12-FEB-1999 (first entry)
XX
DE EST clone GA64.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN WO9845437-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US006956.
XX
PR 10-APR-1997; 97US-00837312.
XX
PA (GEMY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX
XX WPI: 1999-070078/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from e.g.
XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,
XX pituitary, retina and colon cDNA libraries.
XX
XX Claim 1; Page 304; 641pp; English.
XX
XX The present sequence represents an expressed sequence tag (EST), and is a
XX polynucleotide of the invention. The polynucleotides of the invention are
XX all secreted EST sequences isolated from a variety of human tissue
XX sources. The EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX activity. The EST sequences are also stated to be useful for gene therapy
XX
XX Sequence 305 BP; 89 A; 63 C; 77 G; 76 T; 0 U; 0 Other;
XX
XX Query Match 5.7%; Score 222; DB 2; Length 305;
XX Best Local Similarity 99.6%; Pred. No. 2.7e-78;
XX Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 2751 GTTAGATGAAATCTAAGGATGAATACCTTTCAGGAAATCATGGAAGAAATATGAAGATAT 2810
XX 22 GTTAGATGAAATCTAAGGATGAATACCTTTCAGGAAATCATGGAAGAAATATGAAGATAT 81
XX 2811 TAGACAGGACCAATTATGAGTCTCTCAAGGAGAGGAGATATTAACCTTAAGTCAAGCCAG 2870
XX 82 TAGACAGGAGCCATTATGAGTCTCTCAAGGAGAGGAGATATTAACCTTAAGTCAAGCCAG 141
XX 2871 AAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCAACCCAGTGAAGCCACGTTTAT 2930
XX 142 AAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCAACCCAGTGAAGCCACGTTTAT 201
XX 2931 TGGGACCCAGGCTTTTGAAGACTATGACCTGAGAGCTGGTGGACTACATTGACTGGAA 2990
XX 202 TGGGACCCAGGCTTTTGAAGACTATGACCTGAGAGCTGGTGGACTACATTGACTGGAA 261
XX 2991 GCCTTTCTTTGATGTCTGGCAGCTCCGGGGCAA 3023
XX 262 GCCTTTCTTTGATGTCTGGCAGCTCCGGGGCAA 294
XX
```





```
DE XX Human single nucleotide polymorphism (SNP) MTR 5.
KW KW Human; resequence; genotype; disease; forensic; paternity testing;
XX KW single nucleotide polymorphism; SNP; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT variation 16
XX FT /*tag= a
XX FT /standard_name= "single nucleotide polymorphism"
XX PN WO200166800-A2.
XX PD 13-SEP-2001.
XX PF 07-MAR-2001; 2001WO-US007268.
XX PR 07-MAR-2000; 2000US-0187510P.
XX PR 22-MAY-2000; 2000US-0206129P.
XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX PI Cargill M, Ireland JS, Lander ES;
XX XX WPI; 2001-522952/57.
XX PT Nucleic acid molecules from the human genome which include polymorphic
XX PT sites, useful in methods for predicting the presence, absence or severity
XX PT of a particular phenotype or disorder (e.g. diabetes) associated with a
XX PT particular genotype.
XX PS Claim 1; Page 61; 145pp; English.
XX CC The invention relates to the identification of nucleic acid molecules
XX CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
XX CC which can predispose individuals to disease. Various genes from a number
XX CC of individuals were resequenced and single nucleotide polymorphisms
XX CC (SNPs) in these genes discovered. The method is useful for predicting the
XX CC presence, absence or severity of a particular phenotype or disorder (e.g.
XX CC diabetes) associated with a particular genotype. The nucleic acids
XX CC containing the polymorphic sites may be useful in forensics and paternity
XX CC testing
XX CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key
XX SQ Sequence 31 BP; 8 A; 8 C; 8 G; 7 T; 0 U; 0 Other;
XX Query Match 0.8%; Score 31; DB 4; Length 31;
XX Best Local Similarity 100.0%; Pred. NO. 0.068;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 359 GCAGCAGTACTATGCGCCAGCTGACTATGG 389
XX Db 1 GCAGCAGTACTATGCGCCAGCTGACTATGG 31
XX RESULT 12
XX AAI29985
XX ID AAI29985 standard; DNA; 31 BP.
XX AC AAI29985;
XX DT 04-NOV-2004 (revised)
XX DT 18-OCT-2001 (first entry)
XX DE Human single nucleotide polymorphism (SNP) MTR 1.
XX KW Human; resequence; genotype; disease; forensic; paternity testing;
XX KW single nucleotide polymorphism; SNP; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT variation 16
XX FT /*tag= a
XX FT /standard_name= "single nucleotide polymorphism"
XX PN WO200166800-A2.
XX PD 13-SEP-2001.
XX PF 07-MAR-2001; 2001WO-US007268.
XX PR 07-MAR-2000; 2000US-0187510P.
XX PR 22-MAY-2000; 2000US-0206129P.
XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX PI Cargill M, Ireland JS, Lander ES;
XX XX WPI; 2001-522952/57.
XX PT Nucleic acid molecules from the human genome which include polymorphic
XX PT sites, useful in methods for predicting the presence, absence or severity
XX PT of a particular phenotype or disorder (e.g. diabetes) associated with a
XX PT particular genotype.
XX PS Claim 1; Page 61; 145pp; English.
XX CC The invention relates to the identification of nucleic acid molecules
XX CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
XX CC which can predispose individuals to disease. Various genes from a number
XX CC of individuals were resequenced and single nucleotide polymorphisms
XX CC (SNPs) in these genes discovered. The method is useful for predicting the
XX CC presence, absence or severity of a particular phenotype or disorder (e.g.
XX CC diabetes) associated with a particular genotype. The nucleic acids
XX CC containing the polymorphic sites may be useful in forensics and paternity
XX CC testing
XX CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key
XX SQ Sequence 31 BP; 8 A; 8 C; 8 G; 7 T; 0 U; 0 Other;
XX Query Match 0.8%; Score 31; DB 4; Length 31;
XX Best Local Similarity 100.0%; Pred. NO. 0.068;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 359 GCAGCAGTACTATGCGCCAGCTGACTATGG 389
XX Db 1 GCAGCAGTACTATGCGCCAGCTGACTATGG 31
XX RESULT 13
XX AAI29988
XX ID AAI29988 standard; DNA; 31 BP.
XX AC AAI29988;
XX DT 04-NOV-2004 (revised)
XX DT 18-OCT-2001 (first entry)
XX DE Human single nucleotide polymorphism (SNP) MTR 4.
XX KW Human; resequence; genotype; disease; forensic; paternity testing;
XX KW single nucleotide polymorphism; SNP; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT variation 16
XX FT /*tag= a
XX FT /standard_name= "single nucleotide polymorphism"
XX PN WO200166800-A2.
XX PD 13-SEP-2001.
XX PF 07-MAR-2001; 2001WO-US007268.
XX PR 07-MAR-2000; 2000US-0187510P.
XX PR 22-MAY-2000; 2000US-0206129P.
XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX PI Cargill M, Ireland JS, Lander ES;
XX XX WPI; 2001-522952/57.
XX PT Nucleic acid molecules from the human genome which include polymorphic
XX PT sites, useful in methods for predicting the presence, absence or severity
XX PT of a particular phenotype or disorder (e.g. diabetes) associated with a
XX PT particular genotype.
XX PS Claim 1; Page 61; 145pp; English.
XX CC The invention relates to the identification of nucleic acid molecules
XX CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
XX CC which can predispose individuals to disease. Various genes from a number
XX CC of individuals were resequenced and single nucleotide polymorphisms
XX CC (SNPs) in these genes discovered. The method is useful for predicting the
XX CC presence, absence or severity of a particular phenotype or disorder (e.g.
XX CC diabetes) associated with a particular genotype. The nucleic acids
XX CC containing the polymorphic sites may be useful in forensics and paternity
XX CC testing
XX CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key
XX SQ Sequence 31 BP; 5 A; 11 C; 8 G; 7 T; 0 U; 0 Other;
XX Query Match 0.8%; Score 31; DB 4; Length 31;
XX Best Local Similarity 100.0%; Pred. NO. 0.068;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 3192 CATTACCTGTACGAGAGGCTGCTGTGCC 3222
XX Db 1 CATTACCTGTACGAGAGGCTGCTGTGCC 31
XX RESULT 13
XX AAI29988
XX ID AAI29988 standard; DNA; 31 BP.
XX AC AAI29988;
XX DT 04-NOV-2004 (revised)
XX DT 18-OCT-2001 (first entry)
XX DE Human single nucleotide polymorphism (SNP) MTR 4.
XX KW Human; resequence; genotype; disease; forensic; paternity testing;
XX KW single nucleotide polymorphism; SNP; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT variation 16
XX FT /*tag= a
XX FT /standard_name= "single nucleotide polymorphism"
XX PN WO200166800-A2.
XX PD 13-SEP-2001.
XX PF 07-MAR-2001; 2001WO-US007268.
XX PR 07-MAR-2000; 2000US-0187510P.
XX PR 22-MAY-2000; 2000US-0206129P.
XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX PI Cargill M, Ireland JS, Lander ES;
XX XX WPI; 2001-522952/57.
XX PT Nucleic acid molecules from the human genome which include polymorphic
XX PT sites, useful in methods for predicting the presence, absence or severity
XX PT of a particular phenotype or disorder (e.g. diabetes) associated with a
XX PT particular genotype.
XX PS Claim 1; Page 61; 145pp; English.
XX CC The invention relates to the identification of nucleic acid molecules
XX CC (AAI29513-AAI31314) from the human genome which include polymorphic sites
XX CC which can predispose individuals to disease. Various genes from a number
XX CC of individuals were resequenced and single nucleotide polymorphisms
XX CC (SNPs) in these genes discovered. The method is useful for predicting the
XX CC presence, absence or severity of a particular phenotype or disorder (e.g.
XX CC diabetes) associated with a particular genotype. The nucleic acids
XX CC containing the polymorphic sites may be useful in forensics and paternity
XX CC testing
XX CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key
XX SQ Sequence 31 BP; 5 A; 11 C; 8 G; 7 T; 0 U; 0 Other;
XX Query Match 0.8%; Score 31; DB 4; Length 31;
XX Best Local Similarity 100.0%; Pred. NO. 0.068;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 3192 CATTACCTGTACGAGAGGCTGCTGTGCC 3222
XX Db 1 CATTACCTGTACGAGAGGCTGCTGTGCC 31
```

PD 13-SEP-2001.  
 XX 07-MAR-2001; 2001WO-US007268.  
 XX  
 XX 07-MAR-2000; 2000US-0187510P.  
 PR 22-MAY-2000; 2000US-0206129P.  
 XX  
 XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 PA Cargill M, Ireland JS, Lander ES;  
 XX WPI; 2001-522952/57.  
 XX  
 XX Nucleic acid molecules from the human genome which include polymorphic  
 XX sites, useful in methods for predicting the presence, absence or severity  
 XX of a particular phenotype or disorder (e.g. diabetes) associated with a  
 XX particular genotype.  
 XX  
 XX Claim 1; Page 61; 145pp; English.  
 XX  
 XX The invention relates to the identification of nucleic acid molecules  
 XX (AA129513-AA131314) from the human genome which include polymorphic sites  
 XX which can predispose individuals to disease. Various genes from a number  
 XX of individuals were resequenced and single nucleotide polymorphisms  
 XX (SNPs) in these genes discovered. The method is useful for predicting the  
 XX presence, absence or severity of a particular phenotype or disorder (e.g.  
 XX diabetes) associated with a particular genotype. The nucleic acids  
 XX containing the polymorphic sites may be useful in forensics and paternity  
 XX testing  
 XX  
 XX Revised record issued on 04-NOV-2004 : Correction to Feature Table Key  
 XX Sequence 31 BP; 11 A; 7 C; 4 G; 9 T; 0 U; 0 Other;  
 XX  
 XX Query Match 0.8%; Score 31; DB 4; Length 31;  
 XX Best Local Similarity 100.0%; Pred. No. 0.068;  
 XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 348 AATACCTTTAGCAGCAGCTAGTATTGCCAA 378  
 Db 1 AATACCTTTAGCAGCAGCTAGTATTGCCAA 31  
 XX  
 RESULT 14  
 AA129986  
 ID AAT129986 standard; DNA; 31 BP.  
 AC AA129986;  
 XX  
 DT 04-NOV-2004 (revised)  
 DT 18-OCT-2001 (first entry)  
 XX  
 DE Human single nucleotide polymorphism (SNP) MTR 2.  
 XX  
 XX Human; resequence; genotype; disease; forensic; paternity testing;  
 XX single nucleotide polymorphism; SNP; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX variation 16  
 XX /\*tag= a  
 XX /standard\_name= "single nucleotide polymorphism"  
 XX  
 XX WO200166800-A2.  
 XX  
 XX 13-SEP-2001.  
 XX  
 XX 07-MAR-2001; 2001WO-US007268.  
 XX  
 XX 07-MAR-2000; 2000US-0187510P.  
 PR 22-MAY-2000; 2000US-0206129P.  
 XX  
 XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 PA Cargill M, Ireland JS, Lander ES;  
 XX WPI; 2001-522952/57.  
 XX  
 XX Nucleic acid molecules from the human genome which include polymorphic

PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX Cargill M, Ireland JS, Lander ES;  
 XX WPI; 2001-522952/57.  
 XX  
 XX Nucleic acid molecules from the human genome which include polymorphic  
 XX sites, useful in methods for predicting the presence, absence or severity  
 XX of a particular phenotype or disorder (e.g. diabetes) associated with a  
 XX particular genotype.  
 XX  
 XX Claim 1; Page 61; 145pp; English.  
 XX  
 XX The invention relates to the identification of nucleic acid molecules  
 XX (AA129513-AA131314) from the human genome which include polymorphic sites  
 XX which can predispose individuals to disease. Various genes from a number  
 XX of individuals were resequenced and single nucleotide polymorphisms  
 XX (SNPs) in these genes discovered. The method is useful for predicting the  
 XX presence, absence or severity of a particular phenotype or disorder (e.g.  
 XX diabetes) associated with a particular genotype. The nucleic acids  
 XX containing the polymorphic sites may be useful in forensics and paternity  
 XX testing  
 XX  
 XX Revised record issued on 04-NOV-2004 : Correction to Feature Table Key  
 XX Sequence 31 BP; 8 A; 5 C; 10 G; 8 T; 0 U; 0 Other;  
 XX  
 XX Query Match 0.8%; Score 31; DB 4; Length 31;  
 XX Best Local Similarity 100.0%; Pred. No. 0.068;  
 XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 311 AGGAATACCTGCTGGCTGGGCGAGATATCAT 341  
 Db 1 AGGAATACCTGCTGGCTGGGCGAGATATCAT 31  
 XX  
 RESULT 15  
 AA129992  
 ID AAT129992 standard; DNA; 31 BP.  
 AC AA129992;  
 XX  
 DT 04-NOV-2004 (revised)  
 DT 18-OCT-2001 (first entry)  
 XX  
 DE Human single nucleotide polymorphism (SNP) MTR 8.  
 XX  
 XX Human; resequence; genotype; disease; forensic; paternity testing;  
 XX single nucleotide polymorphism; SNP; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX variation 16  
 XX /\*tag= a  
 XX /standard\_name= "single nucleotide polymorphism"  
 XX  
 XX WO200166800-A2.  
 XX  
 XX 13-SEP-2001.  
 XX  
 XX 07-MAR-2001; 2001WO-US007268.  
 XX  
 XX 07-MAR-2000; 2000US-0187510P.  
 PR 22-MAY-2000; 2000US-0206129P.  
 XX  
 XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 PA Cargill M, Ireland JS, Lander ES;  
 XX WPI; 2001-522952/57.  
 XX  
 XX Nucleic acid molecules from the human genome which include polymorphic

PT sites, useful in methods for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype.

PS Claim 1; Page 61; 145pp; English.

XX The invention relates to the identification of nucleic acid molecules (AAI29513-AAI31314) from the human genome which include polymorphic sites which can predispose individuals to disease. Various genes from a number of individuals were resequenced and single nucleotide polymorphisms (SNPs) in these genes discovered. The method is useful for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype. The nucleic acids containing the polymorphic sites may be useful in forensics and paternity testing

CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key

XX Sequence 31 BP; 7 A; 8 C; 11 G; 5 T; 0 U; 0 Other;

Query Match 0.8%; Score 31; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.068;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3540. CGTCGACAGCTCGAGGTTGCGGTACAG 3570

Db 1. CGTCGACAGCTCGAGGTTGCGGTACAG 31

RESULT 16

AAI29987 ID AAI29987 standard; DNA; 31 BP.

XX AAI29987;

XX 04-NOV-2004 (revised)

DT 18-OCT-2001 (first entry)

XX Human single nucleotide polymorphism (SNP) MTR 3.

XX Human; resequence; genotype; disease; forensic; paternity testing; single nucleotide polymorphism; SNP; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX variation 16

FT /\*tag= a

FT /standard\_name= "single nucleotide polymorphism"

XX WO200166800-A2.

XX 13-SEP-2001.

XX 07-MAR-2001; 2001WO-US007268.

XX 07-MAR-2000; 2000US-0187510P.

PR 22-MAY-2000; 2000US-0206129P.

XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.

XX Cargill M, Ireland JS, Lander ES;

XX WPI; 2001-522952/57.

XX Nucleic acid molecules from the human genome which include polymorphic sites, useful in methods for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype.

PS Claim 1; Page 61; 145pp; English.

XX The invention relates to the identification of nucleic acid molecules

CC (AAI29513-AAI31314) from the human genome which include polymorphic sites which can predispose individuals to disease. Various genes from a number of individuals were resequenced and single nucleotide polymorphisms (SNPs) in these genes discovered. The method is useful for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype. The nucleic acids containing the polymorphic sites may be useful in forensics and paternity testing

CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key

XX Sequence 31 BP; 11 A; 5 C; 8 G; 7 T; 0 U; 0 Other;

Query Match 0.8%; Score 31; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.068;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 GGCTGGGGCAGATATCATTTGAAACAATACT 354

Db 1 GGCTGGGGCAGATATCATTTGAAACAATACT 31

RESULT 17

AAI29990 ID AAI29990 standard; DNA; 31 BP.

XX AAI29990;

XX 04-NOV-2004 (revised)

DT 18-OCT-2001 (first entry)

XX Human single nucleotide polymorphism (SNP) MTR 6.

XX Human; resequence; genotype; disease; forensic; paternity testing; single nucleotide polymorphism; SNP; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX variation 16

FT /\*tag= a

FT /standard\_name= "single nucleotide polymorphism"

XX WO200166800-A2.

XX 13-SEP-2001.

XX 07-MAR-2001; 2001WO-US007268.

XX 07-MAR-2000; 2000US-0187510P.

PR 22-MAY-2000; 2000US-0206129P.

XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.

XX Cargill M, Ireland JS, Lander ES;

XX WPI; 2001-522952/57.

XX Nucleic acid molecules from the human genome which include polymorphic sites, useful in methods for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype.

PS Claim 1; Page 61; 145pp; English.

XX The invention relates to the identification of nucleic acid molecules (AAI29513-AAI31314) from the human genome which include polymorphic sites which can predispose individuals to disease. Various genes from a number of individuals were resequenced and single nucleotide polymorphisms (SNPs) in these genes discovered. The method is useful for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype. The nucleic acids containing the polymorphic sites may be useful in forensics and paternity testing

CC testing  
 CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key  
 XX Sequence 31 BP; 8 A; 3 C; 9 G; 11 T; 0 U; 0 Other;  
 SQ Query Match 0.8%; Score 31; DB 4; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.068;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 988 AAGGATTTTCTATGGATGCTTGGTCAATA 1018  
 DB 1 AAGGATTTTCTATGGATGCTTGGTCAATA 31

RESULT 18  
 AAI29991  
 ID AAI29991 standard; DNA; 31 BP.  
 XX AAI29991;  
 AC AAI29991;  
 XX 04-NOV-2004 (revised)  
 DT 18-OCT-2001 (first entry)  
 XX Human single nucleotide polymorphism (SNP) MTR 7.  
 XX Human; resequence; genotype; disease; forensic; paternity testing;  
 KW single nucleotide polymorphism; SNP; ss.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH variation 16  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX WO200166800-A2.  
 XX 13-SEP-2001.  
 XX 07-MAR-2001; 2001WO-US007268.  
 XX 07-MAR-2000; 2000US-0187510P.  
 PR 22-MAY-2000; 2000US-0206129P.  
 XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX Cargill M, Ireland JS, Lander ES;  
 XX WPI; 2001-522952/57.  
 DR Nucleic acid molecules from the human genome which include polymorphic  
 XX sites, useful in methods for predicting the presence, absence or severity  
 FT of a particular phenotype or disorder (e.g. diabetes) associated with a  
 FT particular genotype.  
 XX Claim 1; Page 61; 145pp; English.  
 XX The invention relates to the identification of nucleic acid molecules  
 CC (AAI29513-AAI31314) from the human genome which include polymorphic sites  
 CC which can predispose individuals to disease. Various genes from a number  
 CC of individuals were resequenced and single nucleotide polymorphisms  
 CC (SNPs) in these genes discovered. The method is useful for predicting the  
 CC presence, absence or severity of a particular phenotype or disorder (e.g.  
 CC diabetes) associated with a particular genotype. The nucleic acids  
 CC containing the polymorphic sites may be useful in forensics and paternity  
 CC testing  
 CC Revised record issued on 04-NOV-2004 : Correction to Feature Table Key  
 XX Sequence 31 BP; 7 A; 9 C; 11 G; 4 T; 0 U; 0 Other;  
 SQ Query Match 0.8%; Score 31; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.068;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3525 CAGTGAGCAGCTGGACGCTGCAGACCTGCGA 3555  
 DB 1 CAGTGAGCAGCTGGACGCTGCAGACCTGCGA 31

RESULT 19  
 AAV34083/C  
 ID AAV34083 standard; DNA; 28 BP.  
 XX AAV34083;  
 AC AAV34083;  
 XX 01-MAR-1999 (first entry)  
 DT Methionine synthase primer 1758.  
 XX Methionine synthase primer 1758.  
 DE Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.  
 XX Synthetic.  
 OS  
 XX CA2217153-A.  
 PN 27-MAY-1998.  
 XX 27-NOV-1997; 97CA-02217153.  
 PF 27-NOV-1996; 96US-0031964P.  
 PR 20-JUN-1997; 97US-0050310P.  
 XX (MART-) MARTINEX R & D INC.  
 XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 XX WPI; 1998-569373/49.  
 DR DNA encoding methionine synthase polypeptide - and corresponding  
 XX polypeptide, cells, antibody and therapeutic methods.  
 XX Claim 49; Page 52; 64pp; English.  
 XX This is the nucleotide sequence of oligonucleotide 1758, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2432-2405 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 XX Sequence 28 BP; 10 A; 7 C; 4 G; 7 T; 0 U; 0 Other;  
 SQ Query Match 0.7%; Score 28; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2468 GAGTTATTGATTTAGGAGTCATGACTCC 2495  
 DB 28 GAGTTATTGATTTAGGAGTCATGACTCC 1

RESULT 20  
 ADO57182/C  
 ID ADO57182 standard; DNA; 27 BP.  
 XX

```

AC ADO57182;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human MTR gene PCR primer #30.
XX
KW ss; PCR; primer; human; thrombosis; MTR.
XX
OS Homo sapiens.
XX
PN DE10237073-A1.
XX
PD 15-FEB-2004.
XX
PF 09-AUG-2002; 2002DE-01037073.
XX
PR 09-AUG-2002; 2002DE-01037073.
XX
PA (OGHA-) OGHAM GMBH.
XX
PI Cullen P, Seedorf U;
XX
PW 2004-215770/21.
XX
PT Determining genetic disposition to thrombosis, by testing for presence of
PT at least two allelic polymorphisms associated with increased risk,
PT followed by multifactorial analysis of the results.
XX
PS Example 1; SEQ ID NO 129; 59pp; German.
XX
CC The invention relates to a method of determining the genetic
CC predisposition of a subject for development of thrombosis which comprises
CC examining at least 2 nucleic acid segments from the genome, followed by
CC multifactorial analysis of the results. The nucleic acid segments can
CC contain allelic polymorphisms whose presence is correlated with an
CC increased risk of thrombosis. The method is used to detect polymorphisms,
CC or combinations of them, associated with increased inherited risk of
CC developing thrombosis. The method can detect the synergistic effects of
CC allelic variants when present together, and requires only a single test
CC kit (reducing time and costs). The present sequence represents a
CC thrombosis predisposition PCR primer.
XX
SQ Sequence 27 BP; 7 A; 6 C; 2 G; 12 T; 0 U; 0 Other;

Query Match 0.7%; Score 27; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2759 AAAATCTAAAGGATGAATACCTTTGAGG 2785
DB 27 AAAATCTAAAGGATGAATACCTTTGAGG 1

RESULT 21
AAV34086
ID AAV34086 standard; DNA; 26 BP.
XX
AC AAV34086;
XX
DT 01-MAR-1999 (first entry)
XX
DE Methionine synthase primer 1772.
XX
KW Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;
KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;
KW ss.
XX
OS Synthetic.
XX
PN CA2217153-A.
XX
PD 27-MAY-1998.
XX
PF 27-NOV-1997; 97CA-02217153.
XX
PR 27-NOV-1996; 96US-0031964P.
XX
PW 20-JUN-1997; 97US-0050310P.
XX
PA (MART-) MARTINEX R & D INC.
XX
PI Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX
PW 1998-569373/49.
XX
PT DNA encoding methionine synthase polypeptide - and corresponding

PF 27-NOV-1997; 97CA-02217153.
XX
PR 27-NOV-1996; 96US-0031964P.
XX
PW 20-JUN-1997; 97US-0050310P.
XX
PA (MART-) MARTINEX R & D INC.
XX
PI Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX
PW 1998-569373/49.
XX
PT DNA encoding methionine synthase polypeptide - and corresponding

This is the nucleotide sequence of oligonucleotide 1772, a primer based
on a highly conserved region found in the methionine synthases of
different organisms. It corresponds to nucleotides 2359-2384 of the human
methionine synthase cDNA open reading frame (see AAV34063). 41 Primers
(see AAV34064-104) based on such conserved regions are provided. These
have been used for human and mouse methionine synthase cDNA cloning,
chromosome mapping and in a claimed method for mutation detection.
Mutations in the human methionine synthase gene are associated with
hyperhomocysteinaemia. The invention relates to the diagnosis and
treatment of patients at risk for methionine synthase deficiency, and
associated altered risk for neural tube defects, cardiovascular disease
and colon cancer. Claimed methods for detecting sequence variants involve
SSCP or RFLP polymorphism analysis

Sequence 26 BP; 10 A; 2 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 0.7%; Score 26; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 ATAGGCAAGACATAGTTCGAGTAGT 2447
DB 1 ATAGGCAAGACATAGTTCGAGTAGT 26

RESULT 22
AAV34073
ID AAV34073 standard; DNA; 26 BP.
XX
AC AAV34073;
XX
DT 01-MAR-1999 (first entry)
XX
DE Methionine synthase primer 407A.
XX
KW Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;
KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;
KW ss.
XX
OS Synthetic.
XX
PN CA2217153-A.
XX
PD 27-MAY-1998.
XX
PF 27-NOV-1997; 97CA-02217153.
XX
PR 27-NOV-1996; 96US-0031964P.
XX
PW 20-JUN-1997; 97US-0050310P.
XX
PA (MART-) MARTINEX R & D INC.
XX
PI Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX
PW 1998-569373/49.
XX
PT DNA encoding methionine synthase polypeptide - and corresponding

```

PT polypeptide, cells, antibody and therapeutic methods.  
 PS Claim 49; Page 51; 64pp; English.  
 XX This is the nucleotide sequence of oligonucleotide 407A, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 151-176 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 XX Sequence 26 BP; 9 A; 4 C; 6 G; 7 T; 0 U; 0 Other;  
 SQ Query Match 0.7%; Score 26; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 214 TTCGAGGTCAGGAATTTAAGATCA 239  
 Db 1 TTCGAGGTCAGGAATTTAAGATCA 26  
 RESULT 23  
 AAV34091  
 ID AAV34091 standard; DNA; 26 BP.  
 XX AC AAV34091;  
 XX 01-MAR-1999 (first entry)  
 XX Methionine synthase primer 1796.  
 XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 XX cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 XX ss.  
 XX Synthetic.  
 XX CA2217153-A.  
 XX 27-MAY-1998.  
 XX 27-NOV-1997; 97CA-02217153.  
 XX 27-NOV-1996; 96US-0031964P.  
 XX 20-JUN-1997; 97US-0050310P.  
 XX (MART-) MARTINEX R & D INC.  
 XX Rosen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 XX WPI; 1998-569373/49.  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 XX polypeptide, cells, antibody and therapeutic methods.  
 XX Claim 49; Page 52; 64pp; English.  
 XX This is the nucleotide sequence of oligonucleotide 1796, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2727-2752 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 XX Sequence 26 BP; 9 A; 4 C; 6 G; 7 T; 0 U; 0 Other;  
 SQ Query Match 0.7%; Score 26; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 214 TTCGAGGTCAGGAATTTAAGATCA 239  
 Db 1 TTCGAGGTCAGGAATTTAAGATCA 26  
 RESULT 24  
 AAV34095/c  
 ID AAV34095 standard; DNA; 26 BP.  
 XX AC AAV34095;  
 XX 01-MAR-1999 (first entry)  
 XX Methionine synthase primer 1808.  
 XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 XX cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 XX ss.  
 XX Synthetic.  
 XX CA2217153-A.  
 XX 27-MAY-1998.  
 XX 27-NOV-1997; 97CA-02217153.  
 XX 27-NOV-1996; 96US-0031964P.  
 XX 20-JUN-1997; 97US-0050310P.  
 XX (MART-) MARTINEX R & D INC.  
 XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 XX WPI; 1998-569373/49.  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 XX polypeptide, cells, antibody and therapeutic methods.  
 XX Claim 49; Page 52; 64pp; English.  
 XX This is the nucleotide sequence of oligonucleotide 1808, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2458-2493 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 XX Sequence 26 BP; 12 A; 2 C; 6 G; 6 T; 0 U; 0 Other;  
 SQ Query Match 0.7%; Score 26; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2790 CATGGAAGATATGAAGATATTAGAC 2815  
 Db 1 CATGGAAGATATGAAGATATTAGAC 26  
 RESULT 25  
 AAV34095/c  
 ID AAV34095 standard; DNA; 26 BP.  
 XX AC AAV34095;  
 XX 01-MAR-1999 (first entry)  
 XX Methionine synthase primer 1808.  
 XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 XX cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 XX ss.  
 XX Synthetic.  
 XX CA2217153-A.  
 XX 27-MAY-1998.  
 XX 27-NOV-1997; 97CA-02217153.  
 XX 27-NOV-1996; 96US-0031964P.  
 XX 20-JUN-1997; 97US-0050310P.  
 XX (MART-) MARTINEX R & D INC.  
 XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 XX WPI; 1998-569373/49.  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 XX polypeptide, cells, antibody and therapeutic methods.  
 XX Claim 49; Page 52; 64pp; English.  
 XX This is the nucleotide sequence of oligonucleotide 1808, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2458-2493 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 XX Sequence 26 BP; 7 A; 6 C; 4 G; 9 T; 0 U; 0 Other;  
 SQ Query Match 0.7%; Score 26; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







```

PF 27-NOV-1997; 97CA-02217153.
XX
PR 27-NOV-1996; 96US-0031964P.
PR 20-JUN-1997; 97US-0050310P.
XX
XX (MART-) MARTINEX R & D INC.
XX
PI Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX WPI; 1998-569373/49.
XX
XX DNA encoding methionine synthase polypeptide - and corresponding
XX polypeptide, cells, antibody and therapeutic methods.
XX
XX Claim 49; Page 52; 64pp; English.
XX
XX This is the nucleotide sequence of oligonucleotide 1907B, a primer based
XX on a highly conserved region found in the methionine synthases of
XX different organisms. It corresponds to nucleotides 2057-2033 of the human
XX methionine synthase cDNA open reading frame (see AAV34063). 41 Primers
XX (see AAV34064-104) based on such conserved regions are provided. These
XX have been used for human and mouse methionine synthase cDNA cloning,
XX chromosome mapping and in a claimed method for mutation detection.
XX Mutations in the human methionine synthase gene are associated with
XX hyperhomocysteinaemia. The invention relates to the diagnosis and
XX treatment of patients at risk for methionine synthase deficiency, and
XX associated altered risk for neural tube defects, cardiovascular disease
XX and colon cancer. Claimed methods for detecting sequence variants involve
XX SSCP or RFLP polymorphism analysis
XX
XX Sequence 25 BP; 5 A; 6 C; 5 G; 9 T; 0 U; 0 Other;
XX
Query Match 0.6%; Score 25; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2096 CCCTTGGAAGGGCATTGAAAACA 2120
DB 25 CCCTTGGAAGGGCATTGAAAACA 1
XX
RESULT 30
AAV34100/c
ID AAV34100 standard; DNA; 25 BP.
XX
AC AAV34100;
XX
XX 01-MAR-1999 (first entry)
XX
XX Methionine synthase primer 2606A.
XX
XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;
XX cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;
XX ss.
XX
XX Synthetic.
XX
XX CA2217153-A.
XX
XX 27-MAY-1998.
XX
XX 27-NOV-1997; 97CA-02217153.
XX
XX 27-NOV-1996; 96US-0031964P.
XX
XX 20-JUN-1997; 97US-0050310P.
XX
XX (MART-) MARTINEX R & D INC.
XX
XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX WPI; 1998-569373/49.
XX
XX DNA encoding methionine synthase polypeptide - and corresponding
XX polypeptide, cells, antibody and therapeutic methods.
XX
XX Claim 49; Page 52; 64pp; English.
XX
XX This is the nucleotide sequence of oligonucleotide 1827, a primer based
XX on a highly conserved region found in the methionine synthases of
XX different organisms. It corresponds to nucleotides 1657-1633 of the human
XX methionine synthase cDNA open reading frame (see AAV34063). 41 Primers
XX (see AAV34064-104) based on such conserved regions are provided. These
XX have been used for human and mouse methionine synthase cDNA cloning,
XX chromosome mapping and in a claimed method for mutation detection.
XX Mutations in the human methionine synthase gene are associated with
XX
XX Sequence 25 BP; 8 A; 7 C; 4 G; 6 T; 0 U; 0 Other;
XX
Query Match 0.6%; Score 25; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1117 GAAGGACATATGTTACTGTCTGTC 1141
DB 25 GAAGGACATATGTTACTGTCTGTC 1
XX
RESULT 31
AAV34096/c
ID AAV34096 standard; DNA; 25 BP.
XX
AC AAV34096;
XX
XX 01-MAR-1999 (first entry)
XX
XX Methionine synthase primer 1827.
XX
XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;
XX cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;
XX ss.
XX
XX Synthetic.
XX
XX CA2217153-A.
XX
XX 27-MAY-1998.
XX
XX 27-NOV-1997; 97CA-02217153.
XX
XX 27-NOV-1996; 96US-0031964P.
XX
XX 20-JUN-1997; 97US-0050310P.
XX
XX (MART-) MARTINEX R & D INC.
XX
XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;
XX WPI; 1998-569373/49.
XX
XX DNA encoding methionine synthase polypeptide - and corresponding
XX polypeptide, cells, antibody and therapeutic methods.
XX
XX Claim 49; Page 52; 64pp; English.
XX
XX This is the nucleotide sequence of oligonucleotide 1827, a primer based
XX on a highly conserved region found in the methionine synthases of
XX different organisms. It corresponds to nucleotides 1657-1633 of the human
XX methionine synthase cDNA open reading frame (see AAV34063). 41 Primers
XX (see AAV34064-104) based on such conserved regions are provided. These
XX have been used for human and mouse methionine synthase cDNA cloning,
XX chromosome mapping and in a claimed method for mutation detection.
XX Mutations in the human methionine synthase gene are associated with
XX
XX Sequence 25 BP; 8 A; 7 C; 4 G; 6 T; 0 U; 0 Other;
XX
Query Match 0.6%; Score 25; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1117 GAAGGACATATGTTACTGTCTGTC 1141
DB 25 GAAGGACATATGTTACTGTCTGTC 1
XX

```

CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 25 BP; 5 A; 7 C; 4 G; 9 T; 0 U; 0 Other;

Query Match 0.6%; Score 25; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1696 ACTGGAATGGAGGACACAACTTGT 1720

DB 25 ACTGGAATGGAGGACACAACTTGT 1

# RESULT 32

AAV34103/C  
 ID AAV34103 standard; DNA; 25 BP.

XX AAV34103;

AC (first entry)

DT 01-MAR-1999

DE Methionine synthase primer 2706B.

XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW 88.

XX Synthetic.

XX CA2217153-A.

XX 27-MAY-1998.

XX 27-NOV-1997; 97CA-02217153.

XX 27-NOV-1996; 96US-0031964P.

XX 20-JUN-1997; 97US-0050310P.

XX (MART-) MARTINEX R & D INC.

XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;

XX WPI; 1998-569373/49.

XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.

XX Claim 49; Page 52; 64pp; English.

XX This is the nucleotide sequence of oligonucleotide 2706B, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 3749-3725 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 25 BP; 6 A; 7 C; 4 G; 8 T; 0 U; 0 Other;

# Query Match

Best Local Similarity 0.6%; Score 25; DB 2; Length 25;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3788 CATTGAGGAACATATCTGTGGC 3812

DB 25 CATTGAGGAACATATCTGTGGC 1

# RESULT 33

AAV34087/C  
 ID AAV34087 standard; DNA; 25 BP.

XX AAV34087;

XX 01-MAR-1999 (first entry)

DT Methionine synthase primer 1773.

XX Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW 88.

XX Synthetic.

XX CA2217153-A.

XX 27-MAY-1998.

XX 27-NOV-1997; 97CA-02217153.

XX 27-NOV-1996; 96US-0031964P.

XX 20-JUN-1997; 97US-0050310P.

XX (MART-) MARTINEX R & D INC.

XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;

XX WPI; 1998-569373/49.

XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.

XX Claim 49; Page 52; 64pp; English.

XX This is the nucleotide sequence of oligonucleotide 1773, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2698-2674 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinaemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 25 BP; 8 A; 7 C; 4 G; 6 T; 0 U; 0 Other;

Query Match 0.6%; Score 25; DB 2; Length 25;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2737 GTGTGTTCCAGCTGTAGATGAAA 2761

DB 25 GTGTGTTCCAGCTGTAGATGAAA 1

# RESULT 34

AAV34077

ID AAV34077 standard; DNA; 25 BP.

XX AAV34077;

XX 01-MAR-1999 (first entry)

XX Methionine synthase primer 1406D.  
 DE  
 XX  
 KW Methionine synthase; human; hyperhomocysteinemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.  
 XX  
 OS Synthetic.  
 XX  
 PN CA2217153-A.  
 XX  
 XX  
 PD 27-MAY-1998.  
 XX  
 XX 27-NOV-1997; 97CA-02217153.  
 PF  
 XX 27-NOV-1996; 96US-0031964P.  
 PR  
 XX 20-JUN-1997; 97US-0050310P.  
 PR  
 XX (WART-) MARTINEX R & D INC.  
 PA  
 XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 PI WPI; 1998-569373/49.  
 XX  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.  
 PT  
 XX Claim 49; Page 51; 64pp; English.  
 PS  
 XX This is the nucleotide sequence of oligonucleotide 1406D, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2170-2194 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 CC  
 XX Sequence 25 BP; 7 A; 5 C; 6 G; 7 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.6%; Score 25; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2233 GGAGCTGGAAAAATGTTTCTACCTC 2257  
 DB 1 GGAGCTGGAAAAATGTTTCTACCTC 25  
 RESULT 35  
 AAV34079/c  
 ID AAV34079 standard; DNA; 25 BP.  
 XX  
 AC AAV34079;  
 XX  
 XX 01-MAR-1999 (first entry)  
 DT  
 DE Methionine synthase primer 1707A.  
 XX  
 XX Methionine synthase; human; hyperhomocysteinemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.  
 XX  
 OS Synthetic.  
 XX  
 PN CA2217153-A.  
 XX  
 PD 27-MAY-1998.  
 XX

XX 27-NOV-1997; 97CA-02217153.  
 PF  
 XX 27-NOV-1996; 96US-0031964P.  
 PR  
 XX 20-JUN-1997; 97US-0050310P.  
 PR  
 XX (WART-) MARTINEX R & D INC.  
 PA  
 XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 PI WPI; 1998-569373/49.  
 XX  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.  
 PT  
 XX Claim 49; Page 51; 64pp; English.  
 PS  
 XX This is the nucleotide sequence of oligonucleotide 1707A, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 2129-2105 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis  
 CC  
 XX Sequence 25 BP; 6 A; 4 C; 7 G; 8 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.6%; Score 25; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2168 CCGGACCTCTCAATATTAATTGAAGG 2192  
 DB 25 CCGGACCTCTCAATATTAATTGAAGG 1  
 RESULT 36  
 ABV61053  
 ID ABV61053 standard; cDNA; 237 BP.  
 XX  
 AC ABV61053;  
 XX  
 XX 13-SEP-2002 (first entry)  
 DT  
 DE Human prostate expression marker cDNA 61044.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 XX 23-AUG-2001.  
 PD  
 XX 20-FEB-2001; 2001WO-US005171.  
 PF  
 XX 17-FEB-2000; 2000US-0183319P.  
 PR  
 XX 16-MAR-2000; 2000US-0189862P.  
 PR  
 XX 25-MAY-2000; 2000US-0207454P.  
 PR  
 XX 09-JUN-2000; 2000US-0211314P.  
 PR  
 XX 18-JUL-2000; 2000US-0219007P.  
 PR  
 XX 13-DEC-2000; 2000US-0255281P.  
 PR  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA  
 XX Schlegel R, Endege WO, Monahan JE;  
 PI

XX WPI; 2001-662795/76.  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 XX Claim 1; Page 11601; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 237 BP; 49 A; 45 C; 43 G; 100 T; 0 U; 0 Other;  
 Query Match 0.6%; Score 25; DB 5; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3863 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 3887  
 Db 10 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 34  
 RESULT 37  
 ABL86554/c  
 ID ABL86554 standard; cDNA; 296 BP.  
 XX  
 AC ABL86554;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human ovarian cancer related cDNA clone SEQ ID NO:9532.  
 XX  
 KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200192581-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US017756.  
 XX  
 PR 26-MAY-2000; 2000US-0207484P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Algate PA, Harlocker SL, Jones R;  
 XX  
 DR WPI; 2002-122075/16.  
 XX  
 PT Composition for therapy and diagnosis of ovarian cancer comprising  
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
 PT polypeptide, antibody specific to polypeptide or T cell expressing  
 PT polypeptide.  
 XX  
 PS Claim 1; SEQ ID NO 9532; 489pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising: carriers  
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),

CC or antigen presenting cells that express (II). (I) has cytostatic  
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
 CC detecting ovarian cancer in a patient's biological sample preferably  
 CC serum or ovarian tissue. The method comprises contacting a biological  
 CC sample from a patient with (IV), detecting the amount of polynucleotide  
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 CC value and thereby detecting ovarian cancer in the patient, where the  
 CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 CC useful for stimulating and/or expanding T cells specific for an ovarian  
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
 CC useful in design and preparation of ribosome molecules for inhibiting  
 CC expression of the tumour polypeptides and proteins in tumour cells; and  
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 CC library using well known techniques  
 XX  
 SQ Sequence 296 BP; 132 A; 65 C; 53 G; 46 T; 0 U; 0 Other;  
 Query Match 0.6%; Score 25; DB 6; Length 296;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3863 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 3887  
 Db 288 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 264  
 RESULT 38  
 ACN52851/c  
 ID ACN52851 standard; cDNA; 433 BP.  
 XX  
 AC ACN52851;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Cotton androecium tissue EST Clone ID: LIB3828-018-Q1-N6-F10, SEQ.7632.  
 XX  
 KW Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;  
 KW variety Nu cotton339; library LIB3828; molecular tag; molecular marker;  
 KW genetic mapping; molecular mapping; seed germination; plant growth;  
 KW plant quality; plant yield; plant breeding; tissue printing; ss.  
 XX  
 OS Gossypium hirsutum.  
 XX  
 PN US2004123340-A1.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 12-DEC-2001; 2001US-00021323.  
 XX  
 PR 14-DEC-2000; 2000US-0255619P.  
 XX  
 PA (DEIK/) DEIKMAN J.  
 PA (FENG/) FENG P C C.  
 PA (FINC/) FINCHER K L.  
 PA (ZIEG/) ZIEGLER T E.  
 XX  
 PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;  
 XX  
 DR WPI; 2004-479808/45.  
 XX  
 PT New isolated nucleic acid molecule that encodes a plant protein or its  
 PT fragment, useful for isolating a variety of agronomically significant  
 PT genes associated with plant growth, quality or yield, and as molecular  
 PT tags to map genes.  
 XX  
 PS Claim 1; SEQ ID NO 7632; 34pp; English.  
 XX  
 CC The invention relates to 17880 cotton expressed sequence tags (ESTs;  
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
 CC from primed or non-primed seeds from variety DP50B, mature seeds from  
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeium  
 CC tissue, developing fibres, carpel walls and septa from variety



CC Nucleon33B. The invention also relates to substantially purified  
 CC proteins or their fragments encoded by nucleic acid molecules of the  
 CC invention, and to transformed plants having a nucleic acid construct  
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
 CC molecular tags to isolate genetic regions, to isolate genes, to map  
 CC genes, to determine gene function and to determine whether genes are  
 CC members of a particular gene family. The nucleic acid molecules may be  
 CC used for isolating a variety of agronomically significant genes  
 CC associated with plant growth, quality, yield, and could also serve as  
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
 CC also useful for identifying genes important in initiating and maintaining  
 CC seed germination or that may be used to mitigate stresses encountered  
 CC during seed germination. The ESTs additionally enable the acquisition of  
 CC promoters and cis-regulatory elements which will be useful to express  
 CC agronomically significant genes in these tissues and/or other tissues,  
 CC and also permits the acquisition of molecular markers useful in breeding  
 CC schemes, genetic and molecular mapping, and in cloning of agronomically  
 CC significant genes. The nucleic acid molecules are further useful for  
 CC detecting the presence or quantity of a protein or mRNA and for  
 CC detecting the presence or quantity of a protein by tissue printing. The  
 CC present sequence represents a specifically claimed EST isolated from a  
 CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the US  
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340  
 XX

Sequence 433 BP; 254 A; 58 C; 101 G; 20 T; 0 U; 0 Other;

Query Match 0.6%; Score 25; DB 13; Length 433;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTTCCTTTT 3887  
 DB 188 TTTTGTGCTTTTTCCTTTT 164

RESULT 39  
 ACN52668/c  
 ID ACN52668 standard; cDNA; 475 BP.  
 XX  
 AC ACN52668;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Cotton androecium tissue EST Clone ID: LIB3828-017-Q1-N6-A10, SEQ:7449.  
 XX  
 KW Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;  
 KW variety Nucleon33B; library LIB3828; molecular tag; molecular marker;  
 KW genetic mapping; molecular mapping; seed germination; plant growth;  
 KW plant quality; plant yield; plant breeding; tissue printing; ss.  
 XX  
 OS Gossypium hirsutum.  
 XX  
 PN US2004123340-A1.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 12-DEC-2001; 2001US-00021323.  
 XX  
 PR 14-DEC-2000; 2000US-0255619P.  
 XX  
 XX (DEIK/) DEIKMAN J.  
 PA (FENG/) FENG P C C.  
 PA (FINC/) FINCHER K L.  
 PA (ZIEG/) ZIEGLER T E.  
 XX  
 XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;  
 XX WP1; 2004-479808/45.  
 XX  
 PT New isolated nucleic acid molecule that encodes a plant protein or its  
 PT fragment, useful for isolating a variety of agronomically significant

PT genes associated with plant growth, quality or yield, and as molecular  
 PT tags to map genes.

Claim 1; SEQ ID NO 7449; 34pp; English.

CC The invention relates to 17880 cotton expressed sequence tags (ESTs;  
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated  
 CC from primed or non-primed seeds from variety DP50B, mature seeds from  
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium  
 CC tissue, developing fibres, carpel walls and septa from variety  
 CC Nucleon33B. The invention also relates to substantially purified  
 CC proteins or their fragments encoded by nucleic acid molecules of the  
 CC invention, and to transformed plants having a nucleic acid construct  
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as  
 CC molecular tags to isolate genetic regions, to isolate genes, to map  
 CC genes, to determine gene function and to determine whether genes are  
 CC members of a particular gene family. The nucleic acid molecules may be  
 CC used for isolating a variety of agronomically significant genes  
 CC associated with plant growth, quality, yield, and could also serve as  
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are  
 CC also useful for identifying genes important in initiating and maintaining  
 CC seed germination or that may be used to mitigate stresses encountered  
 CC during seed germination. The ESTs additionally enable the acquisition of  
 CC promoters and cis-regulatory elements which will be useful to express  
 CC agronomically significant genes in these tissues and/or other tissues,  
 CC and also permits the acquisition of molecular markers useful in breeding  
 CC schemes, genetic and molecular mapping, and in cloning of agronomically  
 CC significant genes. The nucleic acid molecules are further useful for  
 CC detecting the expression level or pattern of a protein or mRNA and for  
 CC detecting the presence or quantity of a protein by tissue printing. The  
 CC present sequence represents a specifically claimed EST isolated from a  
 CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the US  
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340  
 XX

Sequence 475 BP; 217 A; 38 C; 201 G; 19 T; 0 U; 0 Other;

Query Match 0.6%; Score 25; DB 13; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTTCCTTTT 3887  
 DB 235 TTTTGTGCTTTTTCCTTTT 211

RESULT 40  
 AAH03248  
 ID AAH03248 standard; cDNA; 789 BP.  
 XX  
 AC AAH03248;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA clone (5'-primer) SEQ ID NO:83.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX

```

PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 1; SEQ ID NO 83; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 789 BP; 237 A; 145 C; 145 G; 258 T; 0 U; 4 Other;
SQ
Query Match 0.6%; Score 25; DB 4; Length 789;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3862 CTTTTTTTTTTTTTTTGCCTTTT 3886
Db 757 CTTTTTTTTTTTTTTTGCCTTTT 781
RESULT 41
AAAG1270/c
ID AAAG1270 standard; DNA; 1180 BP.
XX
XX AAA61270;
AC
XX 18-OCT-2000 (first entry)
DT
XX Human secreted protein gene 11 clone HFPBY77.
DE
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; AIDS; autoimmune disorder; allergy;
XX cardiovascular; viral; bacterial; fungal infection; immunosuppressive;
XX ds.
XX Homo sapiens.
OS
XX WO200029422-A1.
FN
XX 25-MAY-2000.
PD
XX 09-NOV-1999; 99WO-US026409.
XX
XX 12-NOV-1998; 98US-0108207P.
XX
XX

```

---

```

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Ruben SM, Rosen CA, Ebner R, Florence KA, Young PE;
PI Birse CE, Carter KC, Komatsoulis G;
XX
XX WPI; 2000-387729/33.
XX
XX Novel human secreted proteins useful for diagnosing, preventing, treating
PT and ameliorating a medical condition e.g. cardiovascular disease.
PT
XX Claim 1; Page 234-235; 295pp; English.
XX
XX The present sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number and the clone it was derived from
CC are given in the descriptor line. The invention relates to 31 novel genes
CC and their fragments (nucleic acid sequences: AAA61260-A61293; amino acid
CC sequences AAB12301-B12371) which are useful for preventing, treating or
CC ameliorating medical conditions e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 31 polynucleotides, based on which tissues they are most highly expressed
CC in and include products for the diagnosis or treatment of cancer,
CC tumours, AIDS, autoimmune disorders, allergy, cardiovascular disorders,
CC viral, bacterial and fungal infection. The genes are used to generate
CC fusion proteins by linking to the gene a human immunoglobulin portion
CC (AAAG1231) for increasing stability of the fused protein as compared to
CC the secreted protein only
XX
XX Sequence 1180 BP; 308 A; 290 C; 335 G; 247 T; 0 U; 0 Other;
SQ
Query Match 0.6%; Score 25; DB 3; Length 1180;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3863 TTTTTTTTTTTTTTTTGCCTTTT 3887
Db 1167 TTTTTTTTTTTTTTTTGCCTTTT 1143
RESULT 42
ABX10897/c
ID ABX10897 standard; DNA; 2050 BP.
XX
XX ABX10897;
AC
XX 24-APR-2003 (first entry)
DT
XX DNA encoding human orphan chemokine receptor RDC1.
DE
XX Human; neokine family; NEOKINE-1; non-NEOKINE chemokine; chemokine;
XX proliferative disease; cancer; epithelia cancer; liver cancer;
XX secretory gland cancer; bladder cancer; reproductive tract cancer;
XX central nervous system cancer; connective tissue cancer; inflammation;
XX psoriasis; immune rejection; skin graft; kidney transplant; stroke;
XX ischaemia; viral infection; orphan chemokine receptor; RDC1; gene; ds.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
XX CDS 91..1179
XX /*tag= a
XX /product= "RDC-1"
XX /note= "Orphan chemokine receptor"
XX
XX US2002166133-A1.
FN
XX 07-NOV-2002.
PD
XX 27-AUG-2001; 2001US-00940240.
XX
XX 10-FEB-1998; 98US-00023664.
XX
XX 10-FEB-1999; 99US-00248239.
XX

```

```

XX (MILL-) MILLENNIUM PHARM INC.
XX PA
XX PI Barnes TM, Mackay C;
XX PT WPT: 2003-255230/25.
XX DR P-PSDB; ABG74459.
XX DR
XX PT New NEOKINE polypeptides and nucleic acids, useful as modulating agents
XX PT in regulating a variety of cellular processes, in chromosome mapping,
XX PT tissue typing, and in forensic biology.
XX PS Claim 20; Fig 7; 64pp; English.
XX CC
XX CC The invention describes an isolated NEOKINE nucleic acid molecule (I)
XX CC comprising a sequence of 1656, 300, 1372, 237, 1458 or 285 bp given in
XX CC the specification (designated SI-S6, respectively) or their complements.
XX CC The NEOKINE molecules are useful as modulating agents in regulating a
XX CC variety of cellular processes, as primers or hybridisation probes for the
XX CC detection of NEOKINE-encoding nucleic acids, in screening assays, in
XX CC predictive medicine (e.g. diagnostic assays, prognostic assays,
XX CC monitoring clinical trials, and pharmacogenetics), and in a method of
XX CC treatment (therapeutic or prophylactic). The nucleic acids are useful in
XX CC chromosome mapping, tissue typing, and in forensic biology. The proteins
XX CC can be used to treat disorders characterised by insufficient or excessive
XX CC production of non-NEOKINE chemokine or chemokine forms which have
XX CC decreased or aberrant activity compared to wild type chemokines, and to
XX CC screen drugs or compounds which modulate NEOKINE activity. NEOKINE
XX CC modulators are useful for treating and/or preventing proliferative
XX CC diseases such as cancers of the epithelia, liver, secretory glands,
XX CC bladder, reproductive tract, central nervous system, or connective
XX CC tissues, inflammation, psoriasis, or immune rejection following skin
XX CC graft or kidney transplant, and brain inflammation following stroke,
XX CC ischaemia or viral infection. This sequence encodes the human orphan
XX CC chemokine receptor RDC1 that functions as a receptor for NOEKINES
XX SQ Sequence 2050 BP; 435 A; 578 C; 507 G; 530 T; 0 U; 0 Other;

Query Match          0.6%; Score 25; DB 8; Length 2050;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 3864 TTTTNTTTTTTTTTTGCTTTTAA 3888
      |||||||
Db   2050 TTTTNTTTTTTTTTTGCTTTTAA 2026

RESULT 43
ADK66235/c
ID ADK66235 standard; DNA; 2050 BP.
XX AC ADK66235;
XX DT 06-MAY-2004 (first entry)
XX DE Human NEOKINE-1 receptor RDC1 DNA.
XX KW NEOKINE; NEOKINE-associated disorder; genetic disorder; aminoaciduria;
KW cystinosis; CNS disorder; Alzheimer's disease; epilepsy;
KW Parkinson's disease; liver disorder; skeletal muscle disorder;
KW cellular proliferative disorder; cellular differentiative disorder;
KW cancer; hormonal disorder; diabetes; thyroid disorder; immune disorder;
KW inflammatory disorder; rheumatoid arthritis; osteoarthritis; ulcer;
KW cardiovascular disorder; blood vessel disorder; neutrophil disorder;
KW neutropenia; lupus; testicular disorder; mumps; platelet disorder;
KW tissue typing; gene therapy; human; gene; ds.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX CDS 91..1179
FT /tag= a
FT /product= "NEOKINE-1 receptor RDC1 protein"
FT FT

```

XX	US2004019917-A1.
XX	29-JAN-2004.
XX	14-APR-2003; 2003US-00413899.
XX	10-FEB-1998; 98US-00023664.
XX	10-FEB-1999; 99US-00248239.
XX	27-AUG-2001; 2001US-00940240.
XX	(MILL-) MILLENNIUM PHARM INC.
XX	Barnes TM, Mackay C;
XX	WPI; 2004-132242/13.
XX	P-PSDB; ADK66236.
XX	New isolated nucleic acid molecules encoding NEOKINE proteins, useful for
XX	diagnosing and treating kinase-associated disorders, such as diabetes,
XX	Alzheimer's disease, aminoacidurias.
XX	Claim 20; SEQ ID NO 15; 65pp; English.
XX	The invention relates to NEOKINE proteins and their corresponding nucleic
XX	acid sequences. The nucleic acid molecules, polypeptides and antibodies
XX	of the invention are useful for diagnosing and treating NEOKINE-
XX	associated disorders, such as genetic disorders of the membrane transport
XX	CC (aminoacidurias, cystinosis), CNS disorders (Alzheimer's disease,
XX	epilepsy, Parkinson's disease), liver disorders, skeletal muscle
XX	disorders, cellular proliferative and/or differentiative disorders
XX	(cancer), hormonal disorders (diabetes, thyroid disorders), immune and
XX	inflammatory disorders (rheumatoid arthritis, osteoarthritis, ulcer),
XX	cardiovascular disorders, blood vessel disorders, neutrophil disorders
XX	CC (neutropenia, lupus), testicular disorder (mumps) and platelet
XX	disorders. They are also useful in tissue typing, detection assays, and
XX	predictive medicine. The nucleic acids are useful as surrogate markers
XX	CC and also useful in gene therapy. The present sequence is human NEOKINE-1
XX	receptor RDC1 DNA.
XX	Sequence 2050 BP; 435 A; 578 C; 507 G; 530 T; 0 U; 0 Other;
XX	Query Match 0.6%; Score 25; DB 12; Length 2050;
XX	Best Local Similarity 100.0%; Pred. No. 11;
XX	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	3864 TTTTGTGCTTTTGGCTTTTGA 3888
Db	
Db	2050 TTTTGTGCTTTTGGCTTTTGA 2026
RESULT 44	
ABAL9751/c	
ID	ABAL9751 standard; DNA; 14231 BP.
XX	AC
XX	ABAL9751;
XX	
XX	23-JAN-2002 (first entry)
XX	
XX	Human nervous system related polynucleotide SEQ ID NO 12082.
XX	
XX	Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
XX	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
XX	antiparkinsonian; antischlicking; antianemic; antiarthritic; cancer;
XX	antithematic; hepatotropic; cerebroprotective; antiinflammatory;
XX	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX	antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX	neurological disease; infection; immunogenic; gene therapy; vaccine; ds.
OS	Homo sapiens.
XX	
XX	WO200159063-A2.
XX	

PD	16-AUG-2001.		
XX	17-JAN-2001; 2001WO-US0001334.	PR	29-SEP-2000; 2000US-0236368P.
XX	31-JAN-2000; 2000US-0179065P.	PR	29-SEP-2000; 2000US-0236369P.
PR	04-FEB-2000; 2000US-0180628P.	PR	29-SEP-2000; 2000US-0236370P.
PR	24-FEB-2000; 2000US-0184664P.	PR	02-OCT-2000; 2000US-0236802P.
PR	02-MAR-2000; 2000US-0186350P.	PR	02-OCT-2000; 2000US-0237037P.
PR	16-MAR-2000; 2000US-0189874P.	PR	02-OCT-2000; 2000US-0237038P.
PR	17-MAR-2000; 2000US-0190076P.	PR	02-OCT-2000; 2000US-0237039P.
PR	18-APR-2000; 2000US-0198123P.	PR	02-OCT-2000; 2000US-0237040P.
PR	07-MAY-2000; 2000US-0205151P.	PR	13-OCT-2000; 2000US-0239935P.
PR	19-JUN-2000; 2000US-0209467P.	PR	13-OCT-2000; 2000US-0239937P.
PR	28-JUN-2000; 2000US-0214886P.	PR	20-OCT-2000; 2000US-0240960P.
PR	30-JUN-2000; 2000US-0215135P.	PR	20-OCT-2000; 2000US-0241785P.
PR	07-JUL-2000; 2000US-0216647P.	PR	20-OCT-2000; 2000US-0241786P.
PR	07-JUL-2000; 2000US-0216880P.	PR	20-OCT-2000; 2000US-0241787P.
PR	11-JUL-2000; 2000US-0217496P.	PR	20-OCT-2000; 2000US-0241808P.
PR	14-JUL-2000; 2000US-0218290P.	PR	20-OCT-2000; 2000US-0241809P.
PR	26-JUL-2000; 2000US-0220963P.	PR	20-OCT-2000; 2000US-0241826P.
PR	26-JUL-2000; 2000US-0220964P.	PR	20-OCT-2000; 2000US-0242221P.
PR	14-AUG-2000; 2000US-0224518P.	PR	01-NOV-2000; 2000US-0244617P.
PR	14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000; 2000US-0246474P.
PR	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000; 2000US-0246475P.
PR	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000; 2000US-0246476P.
PR	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000; 2000US-0246477P.
PR	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000; 2000US-0246478P.
PR	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000; 2000US-0246523P.
PR	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000; 2000US-0246524P.
PR	14-AUG-2000; 2000US-0225447P.	PR	08-NOV-2000; 2000US-0246525P.
PR	14-AUG-2000; 2000US-0225757P.	PR	08-NOV-2000; 2000US-0246526P.
PR	14-AUG-2000; 2000US-0225758P.	PR	08-NOV-2000; 2000US-0246527P.
PR	14-AUG-2000; 2000US-0225759P.	PR	08-NOV-2000; 2000US-0246528P.
PR	18-AUG-2000; 2000US-0226279P.	PR	08-NOV-2000; 2000US-0246532P.
PR	22-AUG-2000; 2000US-0226681P.	PR	08-NOV-2000; 2000US-0246609P.
PR	22-AUG-2000; 2000US-0226868P.	PR	08-NOV-2000; 2000US-0246610P.
PR	23-AUG-2000; 2000US-0227009P.	PR	08-NOV-2000; 2000US-0246611P.
PR	30-AUG-2000; 2000US-0228924P.	PR	08-NOV-2000; 2000US-0246613P.
PR	01-SEP-2000; 2000US-0229287P.	PR	17-NOV-2000; 2000US-0249207P.
PR	01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000; 2000US-0249208P.
PR	01-SEP-2000; 2000US-0229344P.	PR	17-NOV-2000; 2000US-0249209P.
PR	01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000; 2000US-0249210P.
PR	05-SEP-2000; 2000US-0229509P.	PR	17-NOV-2000; 2000US-0249211P.
PR	05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000; 2000US-0249212P.
PR	06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000; 2000US-0249213P.
PR	06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000; 2000US-0249214P.
PR	08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000; 2000US-0249215P.
PR	08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000; 2000US-0249216P.
PR	08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000; 2000US-0249217P.
PR	08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000; 2000US-0249218P.
PR	08-SEP-2000; 2000US-0231414P.	PR	17-NOV-2000; 2000US-0249244P.
PR	08-SEP-2000; 2000US-0232080P.	PR	17-NOV-2000; 2000US-0249245P.
PR	08-SEP-2000; 2000US-0232081P.	PR	01-DEC-2000; 2000US-0250391P.
PR	12-SEP-2000; 2000US-0231968P.	PR	01-DEC-2000; 2000US-0251160P.
PR	14-SEP-2000; 2000US-0232397P.	PR	05-DEC-2000; 2000US-0251030P.
PR	14-SEP-2000; 2000US-0232398P.	PR	05-DEC-2000; 2000US-0251988P.
PR	14-SEP-2000; 2000US-0232399P.	PR	06-DEC-2000; 2000US-0256719P.
PR	14-SEP-2000; 2000US-0232400P.	PR	06-DEC-2000; 2000US-0251479P.
PR	14-SEP-2000; 2000US-0232401P.	PR	08-DEC-2000; 2000US-0251856P.
PR	14-SEP-2000; 2000US-0233063P.	PR	08-DEC-2000; 2000US-0251869P.
PR	14-SEP-2000; 2000US-0233064P.	PR	08-DEC-2000; 2000US-0251989P.
PR	14-SEP-2000; 2000US-0233065P.	PR	08-DEC-2000; 2000US-0251990P.
PR	21-SEP-2000; 2000US-0234223P.	PR	11-DEC-2000; 2000US-0254097P.
PR	21-SEP-2000; 2000US-0234274P.	PR	05-JAN-2001; 2001US-0259678P.
PR	25-SEP-2000; 2000US-0234997P.	XX	(HUMA-) HUMAN GENOME SCI INC.
PR	25-SEP-2000; 2000US-0234998P.	FA	Rosen CA, Barash SC, Ruben SM;
PR	26-SEP-2000; 2000US-0235484P.	XX	WPI; 2001-541565/60.
PR	27-SEP-2000; 2000US-0235834P.	PI	
PR	27-SEP-2000; 2000US-0235836P.	XX	
PR	29-SEP-2000; 2000US-0236327P.	XX	
PR	29-SEP-2000; 2000US-0236367P.	DR	







FT variation 110101 /tag= ba  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 110228 /tag= bb  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 110420 /tag= bc  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 117399 /tag= bd  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 117478 /tag= be  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 117729 /tag= bf  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 118132 /tag= bg  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 120210 /tag= bh  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 120523 /tag= bi  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 122339 /tag= bj  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation 122339 /tag= bj  
 FT /standard\_name= "Single nucleotide polymorphism"

Query Match 0.6%; Score 25; DB 14; Length 151909;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3862 CTTTTTTTTTTTTTGGCCTTTT 3886  
 DB 54162 CTTTTTTTTTTTTTGGCCTTTT 54138

RESULT 47  
 ADL13693/C  
 ID ADL13693 standard; DNA; 231222 BP.  
 XX AC ADL13693;  
 XX DT 06-MAY-2004 (first entry)  
 XX DE Osteoarthritis-associated polymorphic nucleotide #225.  
 XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
 XX KW joint space narrowing; osteophyte development; joint pain;  
 XX KW osteoarthritis; SNP; single nucleotide polymorphism.  
 XX OS Homo sapiens.  
 XX PN WO2003054166-A2.  
 XX PD 03-JUL-2003.  
 XX PF 19-DEC-2002; 2002WO-US041225.  
 XX PR 20-DEC-2001; 2001US-0342603P.  
 XX PA (INCY-) INCYTE GENOMICS INC.  
 XX PI Jones KA, Schafer A;  
 XX WIPI; 2003-559141/52.

Determining susceptibility of an individual to joint space narrowing.  
 PT osteophyte development and/or joint pain comprises identifying whether  
 PT the individual has at least one polymorphism in a polynucleotide encoding

PT a protein.  
 PS Disclosure; SEQ ID NO. 225; 297pp; English.  
 XX  
 CC The invention relates to a method of determining susceptibility of an  
 CC individual to joint space narrowing and/or osteophyte development and/or  
 CC joint pain comprising identifying whether the individual has at least one  
 CC polymorphism in a polynucleotide encoding at least one of the protein  
 CC listed in the specification. The methods, composition and agent are  
 CC useful for modulating the susceptibility of an individual to joint space  
 CC narrowing and/or osteophyte development and/or joint pain that is  
 CC associated with a disease, preferably osteoarthritis. The cell line and  
 CC the non-human animal are useful for screening for an agent for diagnosing  
 CC an individual having susceptibility to joint space narrowing and/or  
 CC osteophyte development and/or joint pain. This sequence corresponds to  
 CC the polynucleotide encoding a protein listed in the specification. (Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences).

SQ Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;  
 Query Match 0.6%; Score 25; DB 10; Length 231222;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTTTTTTTTTTTGGCCTTTT 3887  
 DB 138768 TTTTTTTTTTTTTTGGCCTTTT 138744

RESULT 48  
 AAV34076/C  
 ID AAV34076 standard; DNA; 24 BP.  
 XX AC AAV34076;  
 XX DT 01-MAR-1999 (first entry)  
 XX DE Methionine synthase primer 1107A.

KW Methionine synthase; human; hyperhomocysteinaemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.

OS Synthetic.  
 XX CA2217153-A.  
 XX PD 27-MAY-1998.  
 XX PF 27-NOV-1997; 97CA-02217153.  
 XX PR 27-NOV-1996; 96US-0031964P.  
 XX PR 20-JUN-1997; 97US-0050310P.  
 XX PA (MART-) MARTINEX R & D INC.

PI Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;  
 XX WIPI; 1998-569373/49.  
 XX DNA encoding methionine synthase polypeptide - and corresponding  
 XX polypeptide, cells, antibody and therapeutic methods.  
 XX Claim 49; Page 51; 64pp; English.

CC This is the nucleotide sequence of oligonucleotide 1107A, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 3856-3833 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,

CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 24 BP; 4 A; 4 C; 6 G; 10 T; 0 U; 0 Other;

Query Match 0.6%; Score 24; DB 2; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3896 GATCCTCAGGAAATACACCTAG 3919

Db 24 GATCCTCAGGAAATACACCTAG 1

RESULT 49

AAV34078

ID AAV34078 standard; DNA; 24 BP.

XX AC AAV34078;

DT 01-MAR-1999 (first entry)

DE Methionine synthase primer 1706A.

XX Methionine synthase; human; hyperhomocysteinemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.

XX OS Synthetic.

XX CA2217153-A.

XX 27-MAY-1998.

XX 27-NOV-1997; 97CA-02217153.

XX 27-NOV-1996; 96US-0031964P.

XX 20-JUN-1997; 97US-0050310P.

XX (MART-) MARTINEX R & D INC.

XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;

XX WPI; 1998-569373/49.

XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.

XX Claim 49; Page 51; 64pp; English.

XX This is the nucleotide sequence of oligonucleotide 1706A, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 1963-1986 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 24 BP; 10 A; 3 C; 8 G; 3 T; 0 U; 0 Other;

Query Match 0.6%; Score 24; DB 2; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 44;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2026 ACAGGAGGGAAGAAATGCTCATTGAG 2049

Db 1 ACAGGAGGGAAGAAATGCTCATTGAG 24

RESULT 50

AAV34088

ID AAV34088 standard; DNA; 24 BP.

XX AC AAV34088;

DT 01-MAR-1999 (first entry)

DE Methionine synthase primer 1774.

XX Methionine synthase; human; hyperhomocysteinemia; neural tube defect;  
 KW cardiovascular disease; colon cancer; diagnosis; therapy; PCR; primer;  
 KW ss.

XX OS Synthetic.

XX CA2217153-A.

XX 27-MAY-1998.

XX 27-NOV-1997; 97CA-02217153.

XX 27-NOV-1996; 96US-0031964P.

XX 20-JUN-1997; 97US-0050310P.

XX (MART-) MARTINEX R & D INC.

XX Rozen R, Leclerc D, Campeau E, Goyette P, Gravel RA;

XX WPI; 1998-569373/49.

XX DNA encoding methionine synthase polypeptide - and corresponding  
 PT polypeptide, cells, antibody and therapeutic methods.

XX Claim 49; Page 52; 64pp; English.

XX This is the nucleotide sequence of oligonucleotide 1774, a primer based  
 CC on a highly conserved region found in the methionine synthases of  
 CC different organisms. It corresponds to nucleotides 3241-3264 of the human  
 CC methionine synthase cDNA open reading frame (see AAV34063). 41 Primers  
 CC (see AAV34064-104) based on such conserved regions are provided. These  
 CC have been used for human and mouse methionine synthase cDNA cloning,  
 CC chromosome mapping and in a claimed method for mutation detection.  
 CC Mutations in the human methionine synthase gene are associated with  
 CC hyperhomocysteinemia. The invention relates to the diagnosis and  
 CC treatment of patients at risk for methionine synthase deficiency, and  
 CC associated altered risk for neural tube defects, cardiovascular disease  
 CC and colon cancer. Claimed methods for detecting sequence variants involve  
 CC SSCP or RFLP polymorphism analysis

XX Sequence 24 BP; 3 A; 11 C; 3 G; 7 T; 0 U; 0 Other;

Query Match 0.6%; Score 24; DB 2; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3304 TGCTCTCAGACTTCATCGCTCCC 3327

Db 1 TGCTCTCAGACTTCATCGCTCCC 24

Search completed: March 7, 2006, 02:53:16  
 Job time : 2110 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 02:09:54 ; Search time 653 Seconds  
(without alignments)  
10668.092 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcactgtgagagcagc.....ctcaagggaatacaacctag 3919

Scoring table:

OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 18

Total number of hits satisfying chosen parameters: 10958

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3919	100.0	3919	3	US-08-980-326-1
2	3601	91.9	3856	3	US-08-980-326-75
3	3517	89.7	7122	3	US-09-318-448-2
4	3517	89.7	7122	3	US-09-347-878-4
5	3517	89.7	7122	3	US-09-577-266-2
6	3466	88.4	7224	3	US-09-347-878-6
7	3109	79.3	7224	3	US-09-962-665-1
8	3109	79.3	7224	3	US-09-963-333-1
9	3109	79.3	7224	3	US-09-962-677-1
10	28	0.7	28	3	US-08-980-326-45
11	26	0.7	26	3	US-08-980-326-35
12	26	0.7	26	3	US-08-980-326-48
13	26	0.7	26	3	US-08-980-326-53
14	26	0.7	26	3	US-08-980-326-57
15	26	0.7	26	3	US-08-980-326-61
16	26	0.7	601	3	US-09-949-016-163933
17	26	0.7	421118	3	US-09-949-016-16297
18	25	0.6	25	3	US-08-980-326-39
19	25	0.6	25	3	US-08-980-326-41
20	25	0.6	25	3	US-08-980-326-49
21	25	0.6	25	3	US-08-980-326-58
22	25	0.6	25	3	US-08-980-326-59
23	25	0.6	25	3	US-08-980-326-60
24	25	0.6	25	3	US-08-980-326-61

22	0.6	601	3	US-09-949-016-58540	Sequence 58540, A	171	22	0.6	112222	3	US-09-949-016-14324	Sequence 14324, A
22	0.6	601	3	US-09-949-016-76668	Sequence 76668, A	172	22	0.6	113186	3	US-09-949-016-17572	Sequence 17572, A
22	0.6	601	3	US-09-949-016-76668	Sequence 76668, A	173	22	0.6	131332	3	US-09-949-016-15535	Sequence 15535, A
22	0.6	601	3	US-09-949-016-77831	Sequence 77831, A	174	22	0.6	131978	3	US-09-949-016-13117	Sequence 13117, A
22	0.6	601	3	US-09-949-016-77832	Sequence 77832, A	175	22	0.6	152582	3	US-09-949-016-12086	Sequence 12086, A
22	0.6	601	3	US-09-949-016-80323	Sequence 80323, A	176	22	0.6	152583	3	US-09-949-016-17390	Sequence 17390, A
22	0.6	601	3	US-09-949-016-89245	Sequence 89245, A	177	22	0.6	152583	3	US-09-949-016-17391	Sequence 17391, A
22	0.6	601	3	US-09-949-016-89246	Sequence 89246, A	178	22	0.6	161124	3	US-09-949-016-11760	Sequence 11760, A
22	0.6	601	3	US-09-949-016-89246	Sequence 89246, A	179	22	0.6	163181	3	US-09-949-016-13730	Sequence 13730, A
22	0.6	601	3	US-09-949-016-90121	Sequence 90121, A	180	22	0.6	164061	3	US-09-949-016-17422	Sequence 17422, A
22	0.6	601	3	US-09-949-016-135415	Sequence 135415, A	181	22	0.6	174639	3	US-09-949-016-16509	Sequence 16509, A
22	0.6	601	3	US-09-949-016-135416	Sequence 135416, A	182	22	0.6	176373	3	US-09-128-155-17	Sequence 17, Appl
22	0.6	601	3	US-09-949-016-141856	Sequence 141856, A	183	22	0.6	202111	3	US-09-949-016-13877	Sequence 13877, A
22	0.6	601	3	US-09-949-016-141857	Sequence 141857, A	184	22	0.6	237510	3	US-09-949-016-14273	Sequence 14273, A
22	0.6	601	3	US-09-949-016-141858	Sequence 141858, A	185	22	0.6	264358	3	US-09-949-016-15725	Sequence 15725, A
22	0.6	601	3	US-09-949-016-141860	Sequence 141860, A	186	22	0.6	301828	3	US-09-949-016-13969	Sequence 13969, A
22	0.6	601	3	US-09-949-016-150316	Sequence 150316, A	187	22	0.6	363032	3	US-09-949-016-12415	Sequence 12415, A
22	0.6	601	3	US-09-949-016-153315	Sequence 153315, A	188	22	0.6	363033	3	US-09-949-016-15754	Sequence 15754, A
22	0.6	601	3	US-09-949-016-153316	Sequence 153316, A	189	21	0.5	21	3	US-08-980-326-43	Sequence 43, Appl
22	0.6	601	3	US-09-949-016-169173	Sequence 169173, A	190	21	0.5	21	3	US-08-980-326-44	Sequence 44, Appl
22	0.6	601	3	US-09-949-016-190626	Sequence 190626, A	191	21	0.5	21	3	US-08-980-326-51	Sequence 51, Appl
22	0.6	601	3	US-09-949-016-198107	Sequence 198107, A	192	21	0.5	21	3	US-09-091-332A-50	Sequence 50, Appl
22	0.6	601	3	US-09-949-016-198108	Sequence 198108, A	193	21	0.5	53	3	US-09-621-976-16091	Sequence 16091, A
22	0.6	601	3	US-09-949-016-200373	Sequence 200373, A	194	21	0.5	302	3	US-09-621-976-644	Sequence 644, Appl
22	0.6	601	3	US-09-949-016-200374	Sequence 200374, A	195	21	0.5	353	3	US-09-949-016-18786	Sequence 18786, A
22	0.6	601	3	US-09-949-016-200641	Sequence 200641, A	196	21	0.5	601	3	US-09-949-016-18787	Sequence 18787, A
22	0.6	601	3	US-09-949-016-200787	Sequence 200787, A	197	21	0.5	601	3	US-09-949-016-21105	Sequence 21105, A
22	0.6	776	3	US-09-372-422A-43	Sequence 43, Appl	198	21	0.5	601	3	US-09-949-016-23642	Sequence 23642, A
22	0.6	1113	3	US-09-857-498A-1	Sequence 1, Appl	199	21	0.5	601	3	US-09-949-016-27839	Sequence 27839, A
22	0.6	1332	3	US-09-010-147B-9	Sequence 9, Appl	200	21	0.5	601	3	US-09-949-016-28678	Sequence 28678, A
22	0.6	5001	3	US-09-949-016-1495	Sequence 1495, Ap	201	21	0.5	601	3	US-09-949-016-30128	Sequence 30128, A
22	0.6	5286	3	US-09-949-016-14866	Sequence 14866, A	202	21	0.5	601	3	US-09-949-016-30129	Sequence 30129, A
22	0.6	6450	3	US-09-041-886-34	Sequence 34, Appl	203	21	0.5	601	3	US-09-949-016-41973	Sequence 41973, A
22	0.6	6450	3	US-08-453-98B-1	Sequence 1, Appl	204	21	0.5	601	3	US-09-949-016-43159	Sequence 43159, A
22	0.6	6450	3	US-10-027-983-3	Sequence 3, Appl	205	21	0.5	601	3	US-09-949-016-43160	Sequence 43160, A
22	0.6	6450	3	US-09-949-016-18	Sequence 18, Appl	206	21	0.5	601	3	US-09-949-016-43388	Sequence 43388, A
22	0.6	6450	3	US-10-052-092-1	Sequence 1, Appl	207	21	0.5	601	3	US-09-949-016-43389	Sequence 43389, A
22	0.6	6450	3	US-10-052-092-7	Sequence 7, Appl	208	21	0.5	601	3	US-09-949-016-43617	Sequence 43617, A
22	0.6	6450	3	US-10-081-563-1	Sequence 1, Appl	209	21	0.5	601	3	US-09-949-016-43618	Sequence 43618, A
22	0.6	9950	3	US-09-949-016-12075	Sequence 12075, A	210	21	0.5	601	3	US-09-949-016-45150	Sequence 45150, A
22	0.6	13182	3	US-09-817-199-3	Sequence 14226, A	211	21	0.5	601	3	US-09-949-016-45151	Sequence 45151, A
22	0.6	15975	3	US-09-949-016-17469	Sequence 3, Appl	212	21	0.5	601	3	US-09-949-016-46400	Sequence 46400, A
22	0.6	18744	3	US-09-949-016-14052	Sequence 17469, A	213	21	0.5	601	3	US-09-949-016-46401	Sequence 46401, A
22	0.6	24150	3	US-09-949-016-12438	Sequence 14052, A	214	21	0.5	601	3	US-09-949-016-50922	Sequence 50922, A
22	0.6	24707	3	US-09-740-027-3	Sequence 12438, A	215	21	0.5	601	3	US-09-949-016-61046	Sequence 61046, A
22	0.6	24707	3	US-10-214-368-3	Sequence 3, Appl	216	21	0.5	601	3	US-09-949-016-61047	Sequence 61047, A
22	0.6	24720	3	US-09-949-016-12341	Sequence 12341, A	217	21	0.5	601	3	US-09-949-016-61048	Sequence 61048, A
22	0.6	24721	3	US-09-949-016-15610	Sequence 15610, A	218	21	0.5	601	3	US-09-949-016-63070	Sequence 63070, A
22	0.6	26334	3	US-09-949-016-17354	Sequence 17354, A	219	21	0.5	601	3	US-09-949-016-65855	Sequence 65855, A
22	0.6	28556	3	US-09-949-016-13064	Sequence 13064, A	220	21	0.5	601	3	US-09-949-016-66433	Sequence 66433, A
22	0.6	32881	3	US-09-949-016-16048	Sequence 13064, A	221	21	0.5	601	3	US-09-949-016-70325	Sequence 70325, A
22	0.6	34408	3	US-09-949-016-14010	Sequence 16048, A	222	21	0.5	601	3	US-09-949-016-70326	Sequence 70326, A
22	0.6	35840	3	US-09-949-016-12156	Sequence 14010, A	223	21	0.5	601	3	US-09-949-016-70327	Sequence 70327, A
22	0.6	35840	3	US-09-949-016-13907	Sequence 12156, A	224	21	0.5	601	3	US-09-949-016-78512	Sequence 78512, A
22	0.6	35840	3	US-09-949-016-13908	Sequence 13907, A	225	21	0.5	601	3	US-09-949-016-78513	Sequence 78513, A
22	0.6	37269	3	US-09-949-016-16672	Sequence 13908, A	226	21	0.5	601	3	US-09-949-016-78514	Sequence 78514, A
22	0.6	40037	3	US-09-949-016-12715	Sequence 16672, A	227	21	0.5	601	3	US-09-949-016-80971	Sequence 80971, A
22	0.6	40408	3	US-09-949-016-16331	Sequence 12715, A	228	21	0.5	601	3	US-09-949-016-80972	Sequence 80972, A
22	0.6	40505	3	US-09-949-016-13439	Sequence 16331, A	229	21	0.5	601	3	US-09-949-016-81007	Sequence 81007, A
22	0.6	44377	3	US-09-949-016-11840	Sequence 13439, A	230	21	0.5	601	3	US-09-949-016-81008	Sequence 81008, A
22	0.6	44378	3	US-09-949-016-15969	Sequence 11840, A	231	21	0.5	601	3	US-09-949-016-86322	Sequence 86322, A
22	0.6	44971	3	US-09-949-016-17049	Sequence 15969, A	232	21	0.5	601	3	US-09-949-016-86323	Sequence 86323, A
22	0.6	54246	3	US-09-949-016-16206	Sequence 17049, A	233	21	0.5	601	3	US-09-949-016-86671	Sequence 86671, A
22	0.6	72843	3	US-09-949-016-12574	Sequence 16206, A	234	21	0.5	601	3	US-09-949-016-86672	Sequence 86672, A
22	0.6	84296	3	US-09-949-016-17375	Sequence 12574, A	235	21	0.5	601	3	US-09-949-016-88336	Sequence 88336, A
22	0.6	90150	3	US-09-949-016-17383	Sequence 17375, A	236	21	0.5	601	3	US-09-949-016-93771	Sequence 93771, A
22	0.6	93920	3	US-09-949-016-12461	Sequence 17383, A	237	21	0.5	601	3	US-09-949-016-94154	Sequence 94154, A
22	0.6	95621	3	US-09-949-016-16853	Sequence 12461, A	238	21	0.5	601	3	US-09-949-016-107611	Sequence 107611, A
22	0.6	102304	3	US-09-949-016-13237	Sequence 16853, A	239	21	0.5	601	3	US-09-949-016-110729	Sequence 110729, A
22	0.6	105413	3	US-09-949-016-12589	Sequence 13237, A	240	21	0.5	601	3	US-09-949-016-110890	Sequence 110890, A
22	0.6	107937	3	US-09-949-016-17192	Sequence 12589, A	241	21	0.5	601	3	US-09-949-016-111212	Sequence 111212, A
22	0.6	112219	3	US-09-949-016-12453	Sequence 3, Appl	242	21	0.5	601	3	US-09-949-016-116570	Sequence 116570, A
22	0.6	112219	3	US-09-949-016-12453	Sequence 12453, A	243	21	0.5	601	3		

244	21	0.5	601	3	US-09-949-016-127373	Sequence 127373,	c 317	21	0.5	3227	3	US-09-919-497-2	Sequence 2, Appli
245	21	0.5	601	3	US-09-949-016-134858	Sequence 134858,	318	21	0.5	3473	3	US-09-949-016-1148	Sequence 1148, Ap
c 246	21	0.5	601	3	US-09-949-016-135114	Sequence 135114,	c 319	21	0.5	3762	3	US-09-252-991A-13448	Sequence 13448, A
c 247	21	0.5	601	3	US-09-949-016-135456	Sequence 135456,	320	21	0.5	3795	3	US-09-252-991A-13693	Sequence 13693, A
c 248	21	0.5	601	3	US-09-949-016-135457	Sequence 135457,	321	21	0.5	4576	2	US-08-832-883-49	Sequence 49, Appl
249	21	0.5	601	3	US-09-949-016-135575	Sequence 135575,	322	21	0.5	4576	2	US-08-832-883-49	Sequence 49, Appl
250	21	0.5	601	3	US-09-949-016-142101	Sequence 142101,	323	21	0.5	6143	2	US-08-612-521-3	Sequence 3, Appli
251	21	0.5	601	3	US-09-949-016-142509	Sequence 142509,	324	21	0.5	6479	3	US-09-949-016-2149	Sequence 2149, Ap
252	21	0.5	601	3	US-09-949-016-142510	Sequence 142510,	c 325	21	0.5	6930	3	US-09-949-016-1063	Sequence 1063, Ap
253	21	0.5	601	3	US-09-949-016-144581	Sequence 144581,	c 326	21	0.5	7008	3	US-09-949-016-16350	Sequence 16350, A
c 254	21	0.5	601	3	US-09-949-016-144581	Sequence 144581,	c 327	21	0.5	8230	3	US-09-949-016-15445	Sequence 15445, A
c 255	21	0.5	601	3	US-09-949-016-147733	Sequence 147733,	c 328	21	0.5	11378	3	US-09-949-016-13815	Sequence 13815, A
c 256	21	0.5	601	3	US-09-949-016-147734	Sequence 147734,	c 329	21	0.5	12437	3	US-09-949-016-16178	Sequence 16178, A
c 257	21	0.5	601	3	US-09-949-016-147735	Sequence 147735,	c 330	21	0.5	12779	3	US-09-949-016-13081	Sequence 13081, A
258	21	0.5	601	3	US-09-949-016-149439	Sequence 149439,	c 331	21	0.5	14160	3	US-09-949-016-16894	Sequence 16894, A
259	21	0.5	601	3	US-09-949-016-150641	Sequence 150641,	c 332	21	0.5	15133	3	US-09-949-016-15001	Sequence 15001, A
260	21	0.5	601	3	US-09-949-016-150642	Sequence 150642,	c 333	21	0.5	15273	3	US-09-949-016-12356	Sequence 12356, A
261	21	0.5	601	3	US-09-949-016-150643	Sequence 150643,	c 334	21	0.5	15273	3	US-09-949-016-13341	Sequence 13341, A
262	21	0.5	601	3	US-09-949-016-154758	Sequence 154758,	c 335	21	0.5	15853	3	US-09-949-016-12121	Sequence 12121, A
263	21	0.5	601	3	US-09-949-016-158671	Sequence 158671,	c 336	21	0.5	15853	3	US-09-949-016-16051	Sequence 16051, A
c 264	21	0.5	601	3	US-09-949-016-160663	Sequence 160663,	c 337	21	0.5	16835	3	US-09-949-016-17309	Sequence 17309, A
c 265	21	0.5	601	3	US-09-949-016-161467	Sequence 161467,	c 338	21	0.5	17415	3	US-09-949-016-13513	Sequence 13513, A
266	21	0.5	601	3	US-09-949-016-164153	Sequence 164153,	c 339	21	0.5	17644	3	US-09-949-016-16342	Sequence 16342, A
267	21	0.5	601	3	US-09-949-016-164154	Sequence 164154,	c 340	21	0.5	17645	3	US-09-949-016-11805	Sequence 11805, A
c 268	21	0.5	601	3	US-09-949-016-164460	Sequence 164460,	c 341	21	0.5	17645	3	US-09-949-016-14188	Sequence 14188, A
c 269	21	0.5	601	3	US-09-949-016-166101	Sequence 166101,	c 342	21	0.5	18200	3	US-09-949-016-13660	Sequence 13660, A
c 270	21	0.5	601	3	US-09-949-016-170021	Sequence 170021,	c 343	21	0.5	18200	3	US-09-949-016-15661	Sequence 15661, A
c 271	21	0.5	601	3	US-09-949-016-170022	Sequence 170022,	c 344	21	0.5	18591	3	US-09-949-016-14719	Sequence 14719, A
c 272	21	0.5	601	3	US-09-949-016-177023	Sequence 177023,	c 345	21	0.5	19719	3	US-09-949-016-15662	Sequence 15662, A
c 273	21	0.5	601	3	US-09-949-016-179094	Sequence 179094,	c 346	21	0.5	19719	3	US-09-949-016-15663	Sequence 15663, A
c 274	21	0.5	601	3	US-09-949-016-182220	Sequence 182220,	c 347	21	0.5	20354	3	US-09-949-016-13219	Sequence 13219, A
275	21	0.5	601	3	US-09-949-016-185193	Sequence 185193,	c 348	21	0.5	20354	3	US-09-949-002-724	Sequence 724, App
c 276	21	0.5	601	3	US-09-949-016-185194	Sequence 185194,	c 349	21	0.5	20301	3	US-09-949-016-11866	Sequence 11866, A
c 277	21	0.5	601	3	US-09-949-016-200653	Sequence 200653,	c 350	21	0.5	21723	3	US-09-949-016-16383	Sequence 16383, A
c 278	21	0.5	601	3	US-09-949-016-200799	Sequence 200799,	c 351	21	0.5	23932	3	US-09-949-016-16099	Sequence 16099, A
c 279	21	0.5	601	3	US-09-949-016-201587	Sequence 201587,	c 352	21	0.5	23932	3	US-09-949-016-13500	Sequence 13500, A
c 280	21	0.5	601	3	US-09-949-016-201588	Sequence 201588,	c 353	21	0.5	24593	3	US-09-949-016-14444	Sequence 14444, A
c 281	21	0.5	601	3	US-09-949-016-204590	Sequence 204590,	c 354	21	0.5	24593	3	US-09-949-016-15059	Sequence 15059, A
c 282	21	0.5	601	3	US-09-949-016-204591	Sequence 204591,	c 355	21	0.5	24971	3	US-09-949-016-14950	Sequence 14950, A
c 283	21	0.5	601	3	US-09-949-001-106	Sequence 106, App	c 356	21	0.5	24971	3	US-09-949-016-16800	Sequence 16800, A
c 284	21	0.5	601	3	US-09-949-001-107	Sequence 107, App	c 357	21	0.5	26896	3	US-09-949-002-643	Sequence 643, App
c 285	21	0.5	601	3	US-09-949-001-108	Sequence 108, App	c 358	21	0.5	30649	3	US-09-949-002-723	Sequence 723, App
c 286	21	0.5	601	3	US-09-949-001-562	Sequence 562, App	c 359	21	0.5	31623	3	US-09-949-016-15945	Sequence 15945, A
c 287	21	0.5	601	3	US-09-949-001-563	Sequence 563, App	c 360	21	0.5	34276	3	US-09-949-016-12263	Sequence 12263, A
c 288	21	0.5	601	3	US-09-949-001-564	Sequence 564, App	c 361	21	0.5	34276	3	US-09-949-016-16103	Sequence 16103, A
c 289	21	0.5	601	3	US-09-949-002-1284	Sequence 1284, Ap	c 362	21	0.5	34278	3	US-09-949-016-17209	Sequence 17209, A
c 290	21	0.5	601	3	US-09-949-002-2663	Sequence 2663, Ap	c 363	21	0.5	35489	3	US-09-949-016-14223	Sequence 14223, A
c 291	21	0.5	601	3	US-09-949-002-2664	Sequence 2664, Ap	c 364	21	0.5	36016	3	US-09-949-016-11998	Sequence 11998, A
c 292	21	0.5	601	3	US-09-949-002-2665	Sequence 2665, Ap	c 365	21	0.5	36302	3	US-09-949-016-13891	Sequence 13891, A
c 293	21	0.5	601	3	US-09-949-002-5300	Sequence 5300, Ap	c 366	21	0.5	36302	3	US-09-949-016-13566	Sequence 13566, A
c 294	21	0.5	601	3	US-09-949-002-5301	Sequence 5301, Ap	c 367	21	0.5	36899	3	US-09-949-016-15757	Sequence 15757, A
c 295	21	0.5	601	3	US-09-949-002-5302	Sequence 5302, Ap	c 368	21	0.5	40323	3	US-09-949-016-16353	Sequence 16353, A
c 296	21	0.5	601	3	US-09-949-002-10416	Sequence 10416, A	c 369	21	0.5	40323	3	US-09-949-016-14233	Sequence 14233, A
c 297	21	0.5	645	3	US-09-522-217-108	Sequence 108, App	c 370	21	0.5	41612	3	US-09-949-016-12769	Sequence 12769, A
c 298	21	0.5	645	3	US-09-323-246-108	Sequence 108, App	c 371	21	0.5	41639	3	US-09-949-016-12773	Sequence 12773, A
c 299	21	0.5	645	3	US-10-295-723-108	Sequence 108, App	c 372	21	0.5	43981	3	US-09-949-016-16294	Sequence 16294, A
c 300	21	0.5	674	3	US-09-620-4058-465	Sequence 465, App	c 373	21	0.5	44064	3	US-09-949-016-12015	Sequence 12015, A
c 301	21	0.5	674	3	US-09-433-8268-465	Sequence 465, App	c 374	21	0.5	44072	3	US-09-949-016-15757	Sequence 15757, A
c 302	21	0.5	674	3	US-09-604-2874-465	Sequence 465, App	c 375	21	0.5	44715	3	US-09-949-016-14999	Sequence 14999, A
c 303	21	0.5	674	3	US-09-634-759-465	Sequence 465, App	c 376	21	0.5	46288	3	US-09-949-016-15000	Sequence 15000, A
c 304	21	0.5	674	3	US-09-590-751A-465	Sequence 465, App	c 377	21	0.5	46311	3	US-09-949-002-585	Sequence 585, App
c 305	21	0.5	674	3	US-09-551-621-465	Sequence 465, App	c 378	21	0.5	46806	3	US-09-949-002-668	Sequence 668, App
c 306	21	0.5	674	3	US-10-076-622-465	Sequence 465, App	c 379	21	0.5	47677	3	US-09-949-002-842	Sequence 842, App
c 307	21	0.5	679	3	US-09-533-559-7211	Sequence 7211, Ap	c 380	21	0.5	48318	3	US-09-949-016-15717	Sequence 15717, A
c 308	21	0.5	729	3	US-09-614-912-67	Sequence 781, Ap	c 381	21	0.5	50263	3	US-09-949-016-13563	Sequence 13563, A
c 309	21	0.5	1239	3	US-09-252-991A-13894	Sequence 67, Appl	c 382	21	0.5	51101	3	US-09-949-016-12859	Sequence 12859, A
c 310	21	0.5	1356	3	US-09-489-847-42	Sequence 13894, A	c 383	21	0.5	51101	3	US-09-949-016-17036	Sequence 17036, A
c 311	21	0.5	1559	3	US-10-104-047-1278	Sequence 42, Appl	c 384	21	0.5	51403	3	US-09-949-016-15057	Sequence 15057, A
c 312	21	0.5	2465	3	US-10-104-047-1278	Sequence 1278, Ap	c 385	21	0.5	54161	3	US-09-949-016-11905	Sequence 11905, A
c 313	21	0.5	2731	3	US-10-104-047-284	Sequence 284, App	c 386	21	0.5	54950	3	US-09-949-016-14383	Sequence 14383, A
c 314	21	0.5	2766	3	US-09-117-447-5	Sequence 5, Appli	c 387	21	0.5	57218	3	US-09-949-016-16512	Sequence 16512, A
c 315	21	0.5	2846	3	US-08-915-795-1	Sequence 1, Appli	c 388	21	0.5	57932	3	US-09-949-002-639	Sequence 619, App
c 316	21	0.5	2846	3	US-09-296-275-1	Sequence 1, Appli	c 389	21	0.5	57932	3	US-09-949-002-822	Sequence 822, App





C 536	20	0.5	601	3	US-09-949-016-44987	Sequence 44987, A	609	20	0.5	601	3	US-09-949-016-98161	Sequence 98161, A
C 537	20	0.5	601	3	US-09-949-016-44988	Sequence 44988, A	610	20	0.5	601	3	US-09-949-016-98427	Sequence 98427, A
C 538	20	0.5	601	3	US-09-949-016-45453	Sequence 45453, A	611	20	0.5	601	3	US-09-949-016-98693	Sequence 98693, A
C 539	20	0.5	601	3	US-09-949-016-46433	Sequence 46433, A	612	20	0.5	601	3	US-09-949-016-98959	Sequence 98959, A
C 540	20	0.5	601	3	US-09-949-016-48162	Sequence 48162, A	613	20	0.5	601	3	US-09-949-016-99225	Sequence 99225, A
C 541	20	0.5	601	3	US-09-949-016-48305	Sequence 48305, A	614	20	0.5	601	3	US-09-949-016-99491	Sequence 99491, A
C 542	20	0.5	601	3	US-09-949-016-48306	Sequence 48306, A	615	20	0.5	601	3	US-09-949-016-99757	Sequence 99757, A
C 543	20	0.5	601	3	US-09-949-016-48370	Sequence 48370, A	616	20	0.5	601	3	US-09-949-016-10023	Sequence 10023, A
C 544	20	0.5	601	3	US-09-949-016-48809	Sequence 48809, A	617	20	0.5	601	3	US-09-949-016-100289	Sequence 100289, A
C 545	20	0.5	601	3	US-09-949-016-48981	Sequence 48981, A	618	20	0.5	601	3	US-09-949-016-100555	Sequence 100555, A
C 546	20	0.5	601	3	US-09-949-016-49153	Sequence 49153, A	619	20	0.5	601	3	US-09-949-016-100821	Sequence 100821, A
C 547	20	0.5	601	3	US-09-949-016-49329	Sequence 49329, A	620	20	0.5	601	3	US-09-949-016-101125	Sequence 101125, A
C 548	20	0.5	601	3	US-09-949-016-52277	Sequence 52277, A	621	20	0.5	601	3	US-09-949-016-101391	Sequence 101391, A
C 549	20	0.5	601	3	US-09-949-016-52820	Sequence 52820, A	622	20	0.5	601	3	US-09-949-016-101657	Sequence 101657, A
C 550	20	0.5	601	3	US-09-949-016-56573	Sequence 56573, A	623	20	0.5	601	3	US-09-949-016-101923	Sequence 101923, A
C 551	20	0.5	601	3	US-09-949-016-56574	Sequence 56574, A	624	20	0.5	601	3	US-09-949-016-102343	Sequence 102343, A
C 552	20	0.5	601	3	US-09-949-016-57321	Sequence 57321, A	625	20	0.5	601	3	US-09-949-016-102483	Sequence 102483, A
C 553	20	0.5	601	3	US-09-949-016-57322	Sequence 57322, A	626	20	0.5	601	3	US-09-949-016-1031976	Sequence 1031976, A
C 554	20	0.5	601	3	US-09-949-016-57323	Sequence 57323, A	627	20	0.5	601	3	US-09-949-016-112621	Sequence 112621, A
C 555	20	0.5	601	3	US-09-949-016-59847	Sequence 59847, A	628	20	0.5	601	3	US-09-949-016-112621	Sequence 112621, A
C 556	20	0.5	601	3	US-09-949-016-59848	Sequence 59848, A	629	20	0.5	601	3	US-09-949-016-121731	Sequence 121731, A
C 557	20	0.5	601	3	US-09-949-016-59888	Sequence 59888, A	630	20	0.5	601	3	US-09-949-016-128452	Sequence 128452, A
C 558	20	0.5	601	3	US-09-949-016-59889	Sequence 59889, A	631	20	0.5	601	3	US-09-949-016-130329	Sequence 130329, A
C 559	20	0.5	601	3	US-09-949-016-60038	Sequence 60038, A	632	20	0.5	601	3	US-09-949-016-131579	Sequence 131579, A
C 560	20	0.5	601	3	US-09-949-016-60039	Sequence 60039, A	633	20	0.5	601	3	US-09-949-016-132268	Sequence 132268, A
C 561	20	0.5	601	3	US-09-949-016-60282	Sequence 60282, A	634	20	0.5	601	3	US-09-949-016-134730	Sequence 134730, A
C 562	20	0.5	601	3	US-09-949-016-61073	Sequence 61073, A	635	20	0.5	601	3	US-09-949-016-135860	Sequence 135860, A
C 563	20	0.5	601	3	US-09-949-016-62802	Sequence 62802, A	636	20	0.5	601	3	US-09-949-016-135977	Sequence 135977, A
C 564	20	0.5	601	3	US-09-949-016-63215	Sequence 63215, A	637	20	0.5	601	3	US-09-949-016-136094	Sequence 136094, A
C 565	20	0.5	601	3	US-09-949-016-63315	Sequence 63315, A	638	20	0.5	601	3	US-09-949-016-136308	Sequence 136308, A
C 566	20	0.5	601	3	US-09-949-016-64102	Sequence 64102, A	639	20	0.5	601	3	US-09-949-016-137921	Sequence 137921, A
C 567	20	0.5	601	3	US-09-949-016-67688	Sequence 67688, A	640	20	0.5	601	3	US-09-949-016-140422	Sequence 140422, A
C 568	20	0.5	601	3	US-09-949-016-71170	Sequence 71170, A	641	20	0.5	601	3	US-09-949-016-142696	Sequence 142696, A
C 569	20	0.5	601	3	US-09-949-016-72462	Sequence 72462, A	642	20	0.5	601	3	US-09-949-016-142697	Sequence 142697, A
C 570	20	0.5	601	3	US-09-949-016-73107	Sequence 73107, A	643	20	0.5	601	3	US-09-949-016-143291	Sequence 143291, A
C 571	20	0.5	601	3	US-09-949-016-73108	Sequence 73108, A	644	20	0.5	601	3	US-09-949-016-143462	Sequence 143462, A
C 572	20	0.5	601	3	US-09-949-016-73181	Sequence 73181, A	645	20	0.5	601	3	US-09-949-016-148948	Sequence 148948, A
C 573	20	0.5	601	3	US-09-949-016-77726	Sequence 77726, A	646	20	0.5	601	3	US-09-949-016-149015	Sequence 149015, A
C 574	20	0.5	601	3	US-09-949-016-77833	Sequence 77833, A	647	20	0.5	601	3	US-09-949-016-149082	Sequence 149082, A
C 575	20	0.5	601	3	US-09-949-016-79112	Sequence 79112, A	648	20	0.5	601	3	US-09-949-016-149149	Sequence 149149, A
C 576	20	0.5	601	3	US-09-949-016-79113	Sequence 79113, A	649	20	0.5	601	3	US-09-949-016-149721	Sequence 149721, A
C 577	20	0.5	601	3	US-09-949-016-79155	Sequence 79155, A	650	20	0.5	601	3	US-09-949-016-149837	Sequence 149837, A
C 578	20	0.5	601	3	US-09-949-016-79156	Sequence 79156, A	651	20	0.5	601	3	US-09-949-016-149839	Sequence 149839, A
C 579	20	0.5	601	3	US-09-949-016-79157	Sequence 79157, A	652	20	0.5	601	3	US-09-949-016-149936	Sequence 149936, A
C 580	20	0.5	601	3	US-09-949-016-79376	Sequence 79376, A	653	20	0.5	601	3	US-09-949-016-149960	Sequence 149960, A
C 581	20	0.5	601	3	US-09-949-016-80690	Sequence 80690, A	654	20	0.5	601	3	US-09-949-016-150016	Sequence 150016, A
C 582	20	0.5	601	3	US-09-949-016-81395	Sequence 81395, A	655	20	0.5	601	3	US-09-949-016-150017	Sequence 150017, A
C 583	20	0.5	601	3	US-09-949-016-83056	Sequence 83056, A	656	20	0.5	601	3	US-09-949-016-150018	Sequence 150018, A
C 584	20	0.5	601	3	US-09-949-016-83355	Sequence 83355, A	657	20	0.5	601	3	US-09-949-016-150076	Sequence 150076, A
C 585	20	0.5	601	3	US-09-949-016-83383	Sequence 83383, A	658	20	0.5	601	3	US-09-949-016-150077	Sequence 150077, A
C 586	20	0.5	601	3	US-09-949-016-83384	Sequence 83384, A	659	20	0.5	601	3	US-09-949-016-154265	Sequence 154265, A
C 587	20	0.5	601	3	US-09-949-016-89396	Sequence 89396, A	660	20	0.5	601	3	US-09-949-016-154266	Sequence 154266, A
C 588	20	0.5	601	3	US-09-949-016-89471	Sequence 89471, A	661	20	0.5	601	3	US-09-949-016-157237	Sequence 157237, A
C 589	20	0.5	601	3	US-09-949-016-89346	Sequence 89346, A	662	20	0.5	601	3	US-09-949-016-157238	Sequence 157238, A
C 590	20	0.5	601	3	US-09-949-016-90006	Sequence 90006, A	663	20	0.5	601	3	US-09-949-016-157980	Sequence 157980, A
C 591	20	0.5	601	3	US-09-949-016-90115	Sequence 90115, A	664	20	0.5	601	3	US-09-949-016-157981	Sequence 157981, A
C 592	20	0.5	601	3	US-09-949-016-90116	Sequence 90116, A	665	20	0.5	601	3	US-09-949-016-162561	Sequence 162561, A
C 593	20	0.5	601	3	US-09-949-016-90337	Sequence 90337, A	666	20	0.5	601	3	US-09-949-016-164521	Sequence 164521, A
C 594	20	0.5	601	3	US-09-949-016-90338	Sequence 90338, A	667	20	0.5	601	3	US-09-949-016-166103	Sequence 166103, A
C 595	20	0.5	601	3	US-09-949-016-90339	Sequence 90339, A	668	20	0.5	601	3	US-09-949-016-169048	Sequence 169048, A
C 596	20	0.5	601	3	US-09-949-016-90340	Sequence 90340, A	669	20	0.5	601	3	US-09-949-016-169049	Sequence 169049, A
C 597	20	0.5	601	3	US-09-949-016-90345	Sequence 90345, A	670	20	0.5	601	3	US-09-949-016-171151	Sequence 171151, A
C 598	20	0.5	601	3	US-09-949-016-90346	Sequence 90346, A	671	20	0.5	601	3	US-09-949-016-171281	Sequence 171281, A
C 599	20	0.5	601	3	US-09-949-016-90347	Sequence 90347, A	672	20	0.5	601	3	US-09-949-016-172075	Sequence 172075, A
C 600	20	0.5	601	3	US-09-949-016-90348	Sequence 90348, A	673	20	0.5	601	3	US-09-949-016-172171	Sequence 172171, A
C 601	20	0.5	601	3	US-09-949-016-92430	Sequence 92430, A	674	20	0.5	601	3	US-09-949-016-174622	Sequence 174622, A
C 602	20	0.5	601	3	US-09-949-016-92430	Sequence 92430, A	675	20	0.5	601	3	US-09-949-016-174623	Sequence 174623, A
C 603	20	0.5	601	3	US-09-949-016-92961	Sequence 92961, A	676	20	0.5	601	3	US-09-949-016-175407	Sequence 175407, A
C 604	20	0.5	601	3	US-09-949-016-96565	Sequence 96565, A	677	20	0.5	601	3	US-09-949-016-175408	Sequence 175408, A
C 605	20	0.5	601	3	US-09-949-016-96831	Sequence 96831, A	678	20	0.5	601	3	US-09-949-016-175409	Sequence 175409, A
C 606	20	0.5	601	3	US-09-949-016-97097	Sequence 97097, A	679	20	0.5	601	3	US-09-949-016-17753	Sequence 17753, A
C 607	20	0.5	601	3	US-09-949-016-97363	Sequence 97363, A	680	20	0.5	601	3	US-09-949-016-17753	Sequence 17753, A
C 608	20	0.5	601	3	US-09-949-016-97629	Sequence 97629, A	681	20	0.5	601	3	US-09-949-016-178492	Sequence 178492, A
C 609	20	0.5	601	3	US-09-949-016-97895	Sequence 97895, A	682	20	0.5	601	3	US-09-949-016-178492	Sequence 178492, A
C 610	20	0.5	601	3	US-09-949-016-97895	Sequence 97895, A	683	20	0.5	601	3	US-09-949-016-178885	Sequence 178885, A

682	20	0.5	601	3	US-09-949-016-179354	Sequence 179354, App	C 755	20	0.5	2231	3	US-09-949-016-4932	Sequence 4932, App
683	20	0.5	601	3	US-09-949-016-180690	Sequence 180690, App	C 756	20	0.5	2231	3	US-09-949-016-4933	Sequence 4933, App
684	20	0.5	601	3	US-09-949-016-182219	Sequence 182219, App	C 757	20	0.5	2280	3	US-08-813-150-1	Sequence 1, Appli
685	20	0.5	601	3	US-09-949-016-182491	Sequence 182491, App	C 758	20	0.5	2280	3	US-09-546-553-1	Sequence 1, Appli
686	20	0.5	601	3	US-09-949-016-182611	Sequence 182611, App	C 759	20	0.5	2280	3	US-10-349-806-1	Sequence 1, Appli
687	20	0.5	601	3	US-09-949-016-183729	Sequence 183729, App	C 760	20	0.5	2442	3	US-09-575-0818-3	Sequence 3, Appli
688	20	0.5	601	3	US-09-949-016-185771	Sequence 185771, App	C 761	20	0.5	2504	3	US-10-104-047-359	Sequence 359, App
689	20	0.5	601	3	US-09-949-016-185980	Sequence 185980, App	C 762	20	0.5	2504	3	US-10-104-047-568	Sequence 568, App
690	20	0.5	601	3	US-09-949-016-185981	Sequence 185981, App	C 763	20	0.5	3701	3	US-09-023-655-309	Sequence 309, App
691	20	0.5	601	3	US-09-949-016-186548	Sequence 186548, App	C 764	20	0.5	3967	3	US-09-672-743-14	Sequence 14, Appli
692	20	0.5	601	3	US-09-949-016-186548	Sequence 186548, App	C 765	20	0.5	7296	3	US-09-949-016-17545	Sequence 17545, A
693	20	0.5	601	3	US-09-949-016-189065	Sequence 189065, App	C 766	20	0.5	7296	3	US-09-949-016-17332	Sequence 17332, A
694	20	0.5	601	3	US-09-949-016-190060	Sequence 190060, App	C 767	20	0.5	7634	3	US-09-949-016-17610	Sequence 17610, A
695	20	0.5	601	3	US-09-949-016-190248	Sequence 190248, App	C 768	20	0.5	7644	3	US-09-949-016-17610	Sequence 17610, A
696	20	0.5	601	3	US-09-949-016-193148	Sequence 193148, App	C 769	20	0.5	7952	3	US-09-949-016-12326	Sequence 12326, A
697	20	0.5	601	3	US-09-949-016-193258	Sequence 193258, App	C 770	20	0.5	7952	3	US-09-949-016-12326	Sequence 12326, A
698	20	0.5	601	3	US-09-949-016-196314	Sequence 196314, App	C 771	20	0.5	8831	3	US-09-949-016-13504	Sequence 13504, A
699	20	0.5	601	3	US-09-949-016-196315	Sequence 196315, App	C 772	20	0.5	9556	3	US-09-949-016-17303	Sequence 17303, A
700	20	0.5	601	3	US-09-949-016-201293	Sequence 201293, App	C 773	20	0.5	9931	3	US-09-949-016-14377	Sequence 14377, A
701	20	0.5	601	3	US-09-949-016-201294	Sequence 201294, App	C 774	20	0.5	10010	3	US-09-949-003-597	Sequence 597, App
702	20	0.5	601	3	US-09-949-016-201337	Sequence 201337, App	C 775	20	0.5	10010	3	US-09-949-003-840	Sequence 840, App
703	20	0.5	601	3	US-09-949-016-201407	Sequence 201407, App	C 776	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
704	20	0.5	601	3	US-09-949-016-201408	Sequence 201408, App	C 777	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
705	20	0.5	601	3	US-09-949-016-201451	Sequence 201451, App	C 778	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
706	20	0.5	601	3	US-09-949-016-203088	Sequence 203088, App	C 779	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
707	20	0.5	601	3	US-09-949-016-203089	Sequence 203089, App	C 780	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
708	20	0.5	601	3	US-09-949-016-203090	Sequence 203090, App	C 781	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
709	20	0.5	601	3	US-09-949-016-205428	Sequence 205428, App	C 782	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
710	20	0.5	601	3	US-09-949-016-205429	Sequence 205429, App	C 783	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
711	20	0.5	601	3	US-09-949-016-205430	Sequence 205430, App	C 784	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
712	20	0.5	601	3	US-09-949-016-205431	Sequence 205431, App	C 785	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
713	20	0.5	601	3	US-09-949-016-205900	Sequence 205900, App	C 786	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
714	20	0.5	601	3	US-09-949-016-206089	Sequence 206089, App	C 787	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
715	20	0.5	601	3	US-09-949-016-206126	Sequence 206126, App	C 788	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
716	20	0.5	601	3	US-09-949-016-206505	Sequence 206505, App	C 789	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
717	20	0.5	601	3	US-09-949-002-1521	Sequence 1521, App	C 790	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
718	20	0.5	601	3	US-09-949-002-1771	Sequence 1771, App	C 791	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
719	20	0.5	601	3	US-09-949-002-1793	Sequence 1793, App	C 792	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
720	20	0.5	601	3	US-09-949-002-2590	Sequence 2590, App	C 793	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
721	20	0.5	601	3	US-09-949-002-2591	Sequence 2591, App	C 794	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
722	20	0.5	601	3	US-09-949-002-2592	Sequence 2592, App	C 795	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
723	20	0.5	601	3	US-09-949-002-4391	Sequence 4391, App	C 796	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
724	20	0.5	601	3	US-09-949-002-4392	Sequence 4392, App	C 797	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
725	20	0.5	601	3	US-09-949-002-4393	Sequence 4393, App	C 798	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
726	20	0.5	601	3	US-09-949-002-5945	Sequence 5945, App	C 799	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
727	20	0.5	601	3	US-09-949-002-5946	Sequence 5946, App	C 800	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
728	20	0.5	601	3	US-09-949-002-6876	Sequence 6876, App	C 801	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
729	20	0.5	601	3	US-09-949-002-9860	Sequence 9860, App	C 802	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
730	20	0.5	601	3	US-09-949-002-9861	Sequence 9861, App	C 803	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
731	20	0.5	601	3	US-09-949-002-9862	Sequence 9862, App	C 804	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
732	20	0.5	601	3	US-09-949-002-10011	Sequence 10011, App	C 805	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
733	20	0.5	601	3	US-09-820-003C-31	Sequence 31, Appli	C 806	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
734	20	0.5	636	3	US-09-931-181-411	Sequence 411, App	C 807	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
735	20	0.5	636	3	US-09-930-444-411	Sequence 411, App	C 808	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
736	20	0.5	636	3	US-09-937-333-411	Sequence 411, App	C 809	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
737	20	0.5	636	3	US-09-932-598-411	Sequence 411, App	C 810	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
738	20	0.5	700	3	US-09-735-271-287	Sequence 287, App	C 811	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
739	20	0.5	751	3	US-09-621-976-3052	Sequence 3052, App	C 812	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
740	20	0.5	817	3	US-09-270-767-12401	Sequence 12401, App	C 813	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
741	20	0.5	924	3	US-09-118-554-62	Sequence 62, Appli	C 814	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
742	20	0.5	924	3	US-09-118-627-62	Sequence 62, Appli	C 815	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
743	20	0.5	924	3	US-09-602-877A-62	Sequence 62, Appli	C 816	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
744	20	0.5	1134	3	US-09-248-335-29	Sequence 29, Appli	C 817	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
745	20	0.5	1134	3	US-09-270-767-13061	Sequence 13061, App	C 818	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
746	20	0.5	1235	3	US-08-672-850-5	Sequence 5, Appli	C 819	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
747	20	0.5	1235	3	US-09-565-177A-5	Sequence 5, Appli	C 820	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
748	20	0.5	1336	3	US-09-257-179-13	Sequence 13, Appli	C 821	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
749	20	0.5	1346	3	US-09-589-733C-9	Sequence 9, Appli	C 822	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
750	20	0.5	1605	3	US-09-270-767-10725	Sequence 9, Appli	C 823	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
751	20	0.5	1697	3	US-09-346-408-5	Sequence 5, Appli	C 824	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
752	20	0.5	2072	3	US-09-786-681A-1	Sequence 1, Appli	C 825	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
753	20	0.5	2209	2	US-08-514-014-1	Sequence 1, Appli	C 826	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A
754	20	0.5	2209	2	US-08-833-823-1	Sequence 1, Appli	C 827	20	0.5	10387	3	US-09-949-016-17414	Sequence 17414, A



```
c 974 20 0.5 76848 3 US-09-949-016-14201 Sequence 14201, A
c 975 20 0.5 77535 3 US-09-949-016-14279 Sequence 14279, A
c 976 20 0.5 77535 3 US-09-949-016-14280 Sequence 14280, A
c 977 20 0.5 77535 3 US-09-949-016-14281 Sequence 14281, A
c 978 20 0.5 77772 3 US-09-949-016-17417 Sequence 17417, A
c 979 20 0.5 77997 3 US-09-949-016-12249 Sequence 12249, A
c 980 20 0.5 78125 3 US-09-949-016-16006 Sequence 16006, A
c 981 20 0.5 78451 3 US-09-949-016-15312 Sequence 15312, A
c 982 20 0.5 78720 3 US-09-949-016-12710 Sequence 12710, A
c 983 20 0.5 78720 3 US-09-949-016-17283 Sequence 17283, A
c 984 20 0.5 80355 3 US-09-949-016-12735 Sequence 12735, A
c 985 20 0.5 80357 3 US-09-949-016-13572 Sequence 13572, A
c 986 20 0.5 81701 3 US-09-949-016-14891 Sequence 14891, A
c 987 20 0.5 81819 3 US-09-949-016-16661 Sequence 16661, A
c 988 20 0.5 81819 3 US-09-949-016-16661 Sequence 16661, A
c 989 20 0.5 81819 3 US-09-949-016-16662 Sequence 16662, A
c 990 20 0.5 81819 3 US-09-949-016-16662 Sequence 16662, A
c 991 20 0.5 82494 3 US-09-949-016-16937 Sequence 16937, A
c 992 20 0.5 82612 3 US-09-949-016-16823 Sequence 16823, A
c 993 20 0.5 83343 3 US-09-949-016-12242 Sequence 12242, A
c 994 20 0.5 83497 3 US-09-949-016-15517 Sequence 15517, A
c 995 20 0.5 84425 3 US-09-949-016-17402 Sequence 17402, A
c 996 20 0.5 84425 3 US-09-949-016-17402 Sequence 17402, A
c 997 20 0.5 84525 3 US-09-949-016-16678 Sequence 16678, A
c 998 20 0.5 84870 3 US-09-949-016-17547 Sequence 17547, A
c 999 20 0.5 85369 3 US-09-949-016-12171 Sequence 12171, A
1000 20 0.5 85675 3 US-09-949-016-12333 Sequence 12333, A
```

## ALIGNMENTS

```
RESULT 1
US-08-980-326-1
; Sequence 1, Application US/08980326
; Patent No. 6703197
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER
; FILE REFERENCE: 50004/002003
; CURRENT APPLICATION NUMBER: US/08/980,326
; CURRENT FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 60/050,310
; EARLIER FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: 60/031,964
; EARLIER FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3919
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(3919)
; OTHER INFORMATION: Entire cloned cDNA encoding wild type methionine
; OTHER INFORMATION: synthase.
US-08-980-326-1

Query Match 100.08; Score 3919; DB 3; Length 3919;
Best Local Similarity 100.08; Pred. No. 0;
Matches 3919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

c 974 1 GGTACCTGTGGAGACAGCTCTTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60
Db 1 GGTACCTGTGGAGACAGCTCTTCTGCGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60
```





OTHER INFORMATION: G;2758 can be C or G.									
US-08-980-326-75									
Query Match 91.9%; Score 3601; DB 3; Length 3856;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 3851; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
Qy	64	ATGTCAACCGCGCTCCAGAACCTGTCCGAACCCGGAAGGCTCTGAAGAAAACCCCTGCGGGAT	123						
Db	1	ATGTCAACCGCGCTCCAGAACCTGTCCGAACCCGGAAGGCTCTGAAGAAAACCCCTGCGGGAT	60						
Qy	124	GAGATCAATGCCATTTCTGCAGAAAGAGATTAATGGTCTGATGGAGGATGGGACCATG	183						
Db	61	GAGATCAATGCCATTTCTGCAGAAAGAGATTAATGGTCTGATGGAGGATGGGACCATG	120						
Qy	184	ATCCAGCGGGAGAGCTTAAACGGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	243						
Db	121	ATCCAGCGGGAGAGCTTAAACGGAACACTTCCGAGGTCAGGAATTTAAAGATCATGCC	180						
Qy	244	AGGCCGCTGAAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTACCA	303						
Db	181	AGGCCGCTGAAAGGCAACATGACATTTTAAAGTATAACTCAGCCTGATGTCAATTACCA	240						
Qy	304	ATCCATAAGGAATPACTTGTGCTGGGCGAGATATCATTTGAAAACAAATACTTTTAGCAGC	363						
Db	241	ATCCATAAGGAATPACTTGTGCTGGGCGAGATATCATTTGAAAACAAATACTTTTAGCAGC	300						
Qy	364	ACTAGTATGCCCAAGCTGACTATGCTTGAACACTTGGCCTACCGGATGAACATGTGC	423						
Db	301	ACTAGTATGCCCAAGCTGACTATGCTTGAACACTTGGCCTACCGGATGAACATGTGC	360						
Qy	424	TCCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAAATTAAGAGG	483						
Db	361	TCCTGAGGAGTGGCCAGAAAAGCTGCCGAGGAGTAACTCTCCAGACAGGAAATTAAGAGG	420						
Qy	484	TTTGTGCGCAGGGCTCTGGGTCGCACTTAATAAGACACTCTCTGTGTCCCATCTGTGGAA	543						
Db	421	TTTGTGCGCAGGGCTCTGGGTCGCACTTAATAAGACACTCTCTGTGTCCCATCTGTGGAA	480						
Qy	544	AGGCCGATTTATAGGAACATCATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCC	603						
Db	481	AGGCCGATTTATAGGAACATCATTTGATGAGCTTGTGAAGCATACCAAGAGCAGGCC	540						
Qy	604	AAAGGACTCTCGATGCGGGGTTGATATCTTACTCATTTGAAACTATTTTGTACTGTCC	663						
Db	541	AAAGGACTCTCGATGCGGGGTTGATATCTTACTCATTTGAAACTATTTTGTACTGTCC	600						
Qy	664	AATGCCAAGGCAGCCTTTTGTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG	723						
Db	601	AATGCCAAGGCAGCCTTTTGTGCACTCCAAATCTTTTTCAGGAGAAATATGCTCCCGG	660						
Qy	724	CTTATCTTTATTTTCAGGAGCATCGTTGATAAAAAGTGGGCGGACTCTTTCCGGACAGCA	783						
Db	661	CTTATCTTTATTTTCAGGAGCATCGTTGATAAAAAGTGGGCGGACTCTTTCCGGACAGCA	720						
Qy	784	GGAGAGGATTTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAAATTTGT	843						
Db	721	GGAGAGGATTTGTTCATCAGCGTGTCTCATGGAGAACCACTCTGCAATGGATTAAATTTGT	780						
Qy	844	GCTTTGGGTGCAGCTGAGATGAGACCTTTTATTTGAAAATAAATTTGAAAATGTACAAAGCC	903						
Db	781	GCTTTGGGTGCAGCTGAGATGAGACCTTTTATTTGAAAATAAATTTGAAAATGTACAAAGCC	840						
Qy	904	TATGTCTCTGTTATCCAAATGAGGCTCTTCCCAACACTTTTGTGTGATATGATGAACG	963						
Db	841	TATGTCTCTGTTATCCAAATGAGGCTCTTCCCAACACTTTTGTGTGATATGATGAACG	900						
Qy	964	CTTCTATGATGGCCAAAGCACTTAAAGGATTTTTCCTATGATGCTTGGTCAATATAGTT	1023						
Db	901	CTTCTATGATGGCCAAAGCACTTAAAGGATTTTTCCTATGATGCTTGGTCAATATAGTT	960						
Qy	1024	GGAGGATGCTGTGGGTCAACCAACAGATCATATCAGGAAAATTTGCTGAAGCTGTCAAAAAT	1083						



Db 961 GGAGGATGCTGTGGGTCAAACACGAGATCATATCAGGGAAATTTGCTGAGCTGTGAAAAAT 1020  
Qy 1084 TGTAAAGCTAGAGTTCCACTGCACTGCTTTTGAAGGACATATGTTACTGTGCTTA 1143  
Db 1021 TGTAAAGCTAGAGTTCCACTGCACTGCTTTTGAAGGACATATGTTACTGTGCTTA 1080  
Qy 1144 GAGCCCTTCAGGATGGAAGCTGACCACTGTTTGAAGCTGGAAGGCTGTGTTATGTT 1203  
Db 1081 GAGCCCTTCAGGATGGAAGCTGACCACTGTTTGAAGCTGGAAGGCTGTGTTATGTT 1140  
Qy 1204 GCAGGATCAAGGAAGTTTGTCTAACTCATATGCGCAGAACTATATGAAGAGCTGTGTGT 1263  
Db 1141 GCAGGATCAAGGAAGTTTGTCTAACTCATATGCGCAGAACTATATGAAGAGCTGTGTGT 1200  
Qy 1264 GTTGCCAAAGTGCAGGTGGAAGTGGAGCCGAGTGTGATGTCAACATGAGATGAGC 1323  
Db 1201 GTTGCCAAAGTGCAGGTGGAAGTGGAGCCGAGTGTGATGTCAACATGAGATGAGC 1260  
Qy 1324 ATGCTAGATGCTCCAGTGCAGTACGACAGATTTTGCAGCTTAATTTGCTCCGAGCCAGAC 1383  
Db 1261 ATGCTAGATGCTCCAGTGCAGTACGACAGATTTTGCAGCTTAATTTGCTCCGAGCCAGAC 1320  
Qy 1384 ATGCTAGATGCTCCAGTGCAGTACGACAGATTTTGCAGCTTAATTTGCTCCGAGCTGTTA 1443  
Db 1321 ATGCTAGATGCTCCAGTGCAGTACGACAGATTTTGCAGCTTAATTTGCTCCGAGCTGTTA 1380  
Qy 1444 AAGTGTGCCAAGGAAGTGCATTTGCAATAGCATTTAGTCTGAAGGAAGGAGGAGCGAC 1503  
Db 1381 AAGTGTGCCAAGGAAGTGCATTTGCAATAGCATTTAGTCTGAAGGAAGGAGGAGCGAC 1440  
Qy 1504 TTCTTGGAGAGCCAGGAGATTTAAAGATGAGGAGTGTGCTATGTTGCTGATGCTTTT 1563  
Db 1441 TTCTTGGAGAGCCAGGAGATTTAAAGATGAGGAGTGTGCTATGTTGCTGATGCTTTT 1500  
Qy 1564 GATGAAGAAGGACAGGCAACAGAGACAGACAAAAATCAGAGTGTGACCCGGGCTAC 1623  
Db 1501 GATGAAGAAGGACAGGCAACAGAGACAGACAAAAATCAGAGTGTGACCCGGGCTAC 1560  
Qy 1624 CATCTGCTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTGACCCCTAATATC 1683  
Db 1561 CATCTGCTGTGAAAAAACTGGGCTTTAATCCAAATGACATTTATTTGACCCCTAATATC 1620  
Qy 1684 CTAAACCATTTGGAGTGGAGTGGAGGAAACACAACTTGTATGCCATTTATTTATTCATGCA 1743  
Db 1621 CTAAACCATTTGGAGTGGAGTGGAGGAAACACAACTTGTATGCCATTTATTTATTCATGCA 1680  
Qy 1744 ACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGATAGTGGAGTCTTTCCAACTTG 1803  
Db 1681 ACAAAAGTCAATTAAGAAACATTTACCTGGAGCCAGATAGTGGAGTCTTTCCAACTTG 1740  
Qy 1804 TCCTTCTCTCCGAGGAATGGAGCCATTCGAGAAAGCAATGATGGGTTTTCCTTTTAC 1863  
Db 1741 TCCTTCTCTCCGAGGAATGGAGCCATTCGAGAAAGCAATGATGGGTTTTCCTTTTAC 1800  
Qy 1864 CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGTAT 1923  
Db 1801 CATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGTAT 1860  
Qy 1924 GATGATATCCATTAAGGAATCTTCGAGCTCTGTGAGATCTCATCTGGAATTAAGACCTT 1983  
Db 1861 GATGATATCCATTAAGGAATCTTCGAGCTCTGTGAGATCTCATCTGGAATTAAGACCTT 1920  
Qy 1984 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAACAGAGGAGGAGAAAGTC 2043  
Db 1921 GAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAACAGAGGAGGAGAAAGTC 1980  
Qy 2044 ATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCTTGTGATGTCCTTTGTG 2103  
Db 1981 ATTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACGCTTGTGATGTCCTTTGTG 2040  
Qy 2104 AAGGCAATTAAGAAACATATTTATGAGGATCTGAGGAGCCAGGTTTAACCAAAAAA 2163  
Db 2041 AAGGCAATTAAGAAACATATTTATGAGGATCTGAGGAGCCAGGTTTAACCAAAAAA 2100

Qy 2164 TATCCCGGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTGGT 2223  
Db 2101 TATCCCGGACCTCTCAATATAATTTGAAGGACCCCTGATGAATGGAATGAAATTTGTGGT 2160  
Qy 2224 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCGGGTATG 2283  
Db 2161 GATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTTATAAGTCAAGCCGGGTATG 2220  
Qy 2284 AAGAGGCTGTTTGGCCACCTTATCCCTTTCATCGAAAAAAGAGAAAGAAACACAGAGTG 2343  
Db 2221 AAGAGGCTGTTTGGCCACCTTATCCCTTTCATCGAAAAAAGAGAAAGAAACACAGAGTG 2280  
Qy 2344 CTTTAAAGGCAAGTGAAGAGGACCTTACCAGGGCACCATCTGCTGGCCACTGTT 2403  
Db 2281 CTTTAAAGGCAAGTGAAGAGGACCTTACCAGGGCACCATCTGCTGGCCACTGTT 2340  
Qy 2404 AAGGGCGAGCTGCACGACATAGGCAAGACATAGTTGGAGTAGTCTTGGCTGCATAT 2463  
Db 2341 AAGGGCGAGCTGCACGACATAGGCAAGACATAGTTGGAGTAGTCTTGGCTGCATAT 2400  
Qy 2464 TTCCGAGTTTATGATTTAGGAGTCACTGACTCCATGTGATAAGATCTGAAAGCTGCTTT 2523  
Db 2401 TTCCGAGTTTATGATTTAGGAGTCACTGACTCCATGTGATAAGATCTGAAAGCTGCTTT 2460  
Qy 2524 GACCAAAAGCGAGATATAATTTGGCTGTGAGACTCATCTCCTTCCCTGGATGAAATG 2583  
Db 2461 GACCAAAAGCGAGATATAATTTGGCTGTGAGACTCATCTCCTTCCCTGGATGAAATG 2520  
Qy 2584 ATTTTGTGTTCCAAAGAAATGGAGATAGCTATAAGGATTCATTTGTTGATGGAGGA 2643  
Db 2521 ATTTTGTGTTCCAAAGAAATGGAGATAGCTATAAGGATTCATTTGTTGATGGAGGA 2580  
Qy 2644 GCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACCTGTA 2703  
Db 2581 GCAACCACTTCAAAAACCCACACAGCAGTTAAATAGCTCCGAGATACAGTGACCTGTA 2640  
Qy 2704 ATCCATGCTCTGAGAGCGCTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAAT 2763  
Db 2641 NNCCATGCTCTGAGAGCGCTCCAAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAAT 2700  
Qy 2764 CTAAGGATGATATCTTTGAGGAAATCATGGAAGATATGAAGATATTAAGACGACAT 2823  
Db 2701 CTAAGGATGATATCTTTGAGGAAATCATGGAAGATATGAAGATATTAAGACGACAT 2760  
Qy 2824 TATGAGTCTCTCAAGGAGAGGAGATCTTACCTTAAAGTCAAGCCAGAAAAAGTGTTC 2883  
Db 2761 TATGAGTCTCTCAAGGAGAGGAGATCTTACCTTAAAGTCAAGCCAGAAAAAGTGTTC 2820  
Qy 2884 CAAATGGATTTGGCTGTCTGAACCTCAACCAGTGAAGCCACCGTTTATTTGGGACCCAGGTC 2943  
Db 2821 CAAATGGATTTGGCTGTCTGAACCTCAACCAGTGAAGCCACCGTTTATTTGGGACCCAGGTC 2880  
Qy 2944 TTTGAAGACTATGACTGACAGAGCTGTGGAGTACATTTGACTGGAGCTTTCTTTGAT 3003  
Db 2881 TTTGAAGACTATGACTGACAGAGCTGTGGAGTACATTTGACTGGAGCTTTCTTTGAT 2940  
Qy 3004 GTCTGAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTAAACGACAA 3063  
Db 2941 GTCTGAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTCCCAAGATATTTAAACGACAA 3000  
Qy 3064 ACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCACAATATGCTGAACACACTG 3123  
Db 3001 ACAGTAGTGGAGAGGCCAGGAAGTCTACGATGATGCCACAATATGCTGAACACACTG 3060  
Qy 3124 ATTAGTCAAAAAGAACTCCGGGCCCGGGTGTGGTGTGGTCTTGGCCAGACAGAGTATC 3183  
Db 3061 ATTAGTCAAAAAGAACTCCGGGCCCGGGTGTGGTGTGGTCTTGGCCAGACAGAGTATC 3120  
Qy 3184 CAAGGCAATTCACCTGTACCGAGGCTGTGCCCCCAGGCTGCTGCCCCCAGGCTGCAGAGCCATAGCC 3243  
Db 3121 CAAGGCAATTCACCTGTACCGAGGCTGTGCCCCCAGGCTGCTGCCCCCAGGCTGCAGAGCCATAGCC 3180

QY 3244 ACTTCTATGGTTAAGCAACAGGCTGAGAAGACTCTCCAGCAAGGAGCCATCTAC 3303  
DB 3181 ACTTCTATGGTTAAGCAACAGGCTGAGAAGACTCTCCAGCAAGGAGCCATCTAC 3240  
QY 3304 TGCCCTCAGACTTCATCGCTCCCTTGGCATCTGGCATCTGGTACTACCTGGGCTGTTT 3363  
DB 3241 TGCCCTCAGACTTCATCGCTCCCTTGGCATCTGGCATCTGGTACTACCTGGGCTGTTT 3300  
QY 3364 GCCGTTGCCCTGCTTTGGGTTAGAGAGCTGAGCAAGGCCCTATCAGGATGATGTTGACGAC 3423  
DB 3301 GCCGTTGCCCTGCTTTGGGTTAGAGAGCTGAGCAAGGCCCTATCAGGATGATGTTGACGAC 3360  
QY 3424 TACAGCAGCATCATGTGCAAGGGCTGGGGACCGGCTGGCAGAGGCTTTGCAAGAGAG 3483  
DB 3361 TACAGCAGCATCATGTGCAAGGGCTGGGGACCGGCTGGCAGAGGCTTTGCAAGAGAG 3420  
QY 3484 CTCATGAAGAAGTTGCCGAGAACTGTGGGCCCTACTGTGGCGATGAGCAGCTGGAGCTC 3543  
DB 3421 CTCATGAAGAAGTTGCCGAGAACTGTGGGCCCTACTGTGGCGATGAGCAGCTGGAGCTC 3480  
QY 3544 GCAGACTGGAAGGTTGCCGTACAGGGCATCCGCCGGCTCCTGGCTACCCAGCCAG 3603  
DB 3481 GCAGACTGGAAGGTTGCCGTACAGGGCATCCGCCGGCTCCTGGCTACCCAGCCAG 3540  
QY 3604 CCGACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC 3663  
DB 3541 CCGACCAACCCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAGGC 3600  
QY 3664 ATTAGGTTAAGCAATCATTTAGCAATGGCACTGCTGCTTCCAGGATCAGGTTGAGGAT 3723  
DB 3601 ATTAGGTTAAGCAATCATTTAGCAATGGCACTGCTGCTTCCAGGATCAGGTTGAGGAT 3660  
QY 3724 TCCAAATTTGAAGTCCAAATATTTTGGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGGAT 3783  
DB 3661 TCCAAATTTGAAGTCCAAATATTTTGGCTGGGGAAGATTTCCAAAGGATCAGGTTGAGGAT 3720  
QY 3784 TATGCAATGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGAGCACTTTTG 3843  
DB 3721 TATGCAATGAGGAAGAACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGAGCACTTTTG 3780  
QY 3844 GGATATGATCAGACTAACTT 3903  
DB 3781 GGATATGATCAGACTAACTT 3840  
QY 3904 AGGAATACACCTAG 3919  
DB 3841 AGGAATACACCTAG 3856

## RESULT 3

US-09-318-448-2  
; Sequence 2, Application US/09318448  
; Patent No. 6210950  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, William G.  
; APPLICANT: Stenroos, Edward S.  
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS  
; FILE REFERENCE: 601-1-057  
; CURRENT APPLICATION NUMBER: US/09/318,448  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 7122  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-318-448-2

Query Match 89.7%; Score 3517; DB 3; Length 7122;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GTCACTGTGAGAGCACTCTCTCTCTGCGGCCCTCTGCGCAAGGAGAGACTCGACA 61  
DB 225 GTCACTGTGAGAGCACTCTCTCTCTGCGGCCCTCTGCGCAAGGAGAGACTCGACA 284  
QY 62 ACATGTCACTGCGGCTCAAGACCTGTCCCAAGGAGGCTCTGAAGAAAACCTCTGGG 121  
DB 285 ACATGTCACTGCGGCTCAAGACCTGTCCCAAGGAGGCTCTGAAGAAAACCTCTGGG 344  
QY 122 ATGAGATCAATGCACTTCTGCAAGAGAGGATATGCTGTGATGAGGAGTGGGACCA 181  
DB 345 ATGAGATCAATGCACTTCTGCAAGAGAGGATATGCTGTGATGAGGAGTGGGACCA 404  
QY 182 TGATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 241  
DB 405 TGATCCAGCGGAGAGAGCTAAACGAGAACACTTCCGAGGTCAGGAATTTAAAGATCATG 464  
QY 242 CAGGCGGCTGAAAGGCAACATGACATTTTAAAGTATAAATCAGCCTGATGATTTACC 301  
DB 465 CAGGCGGCTGAAAGGCAACATGACATTTTAAAGTATAAATCAGCCTGATGATTTACC 524  
QY 302 AAATCCATAGGAATACTTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTAGCA 361  
DB 525 AAATCCATAGGAATACTTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTAGCA 584  
QY 362 GCACTAGTATGCTCCAGGCTGACTATGGCTTTGAAACAATTGGCCTACCGATGAACATGT 421  
DB 585 GCACTAGTATGCTCCAGGCTGACTATGGCTTTGAAACAATTGGCCTACCGATGAACATGT 644  
QY 422 GCTCTCAGAGTGGGCGAGAAAGCTGCGAGGAGGTAATCTCCAGACAGGAATTAAGA 481  
DB 645 GCTCTCAGAGTGGGCGAGAAAGCTGCGAGGAGGTAATCTCTCCAGACAGGAATTAAGA 704  
QY 482 GGTCTGTGGCAGGGCTCTGGGTCGCACTAAATAAGACACCTCTGTGTGCCCATCTGTGG 541  
DB 705 GGTCTGTGGCAGGGCTCTGGGTCGCACTAAATAAGACACCTCTGTGTGCCCATCTGTGG 764  
QY 542 AAAGGCGGATATAGGAACATCACATTTTGAAGAGCTTGTGAAGCATACCAAGAGCAGG 601  
DB 765 AAAGGCGGATATAGGAACATCACATTTTGAAGAGCTTGTGAAGCATACCAAGAGCAGG 824  
QY 602 CCAAGGACTTCTGGATGGGCGGTTGATATCTTACTCATTTGAACTATTTTGTACTG 661  
DB 825 CCAAGGACTTCTGGATGGGCGGTTGATATCTTACTCATTTGAACTATTTTGTACTG 884  
QY 662 CCAATGCCAAGCAGCCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
DB 885 CCAATGCCAAGCAGCCTTGTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 944  
QY 722 GGCCTATCTTTATTTTCCAGGACGATCGTTGATAAAAAGTGGGCGGACTCTTTCCGGAAGA 781  
DB 945 GGCCTATCTTTATTTTCCAGGACGATCGTTGATAAAAAGTGGGCGGACTCTTTCCGGAAGA 1004  
QY 782 CAGGAGGAGATTTGTGATCAGCGTGTCTCATGAGAGAACACCTCTGCTGATTTGATTAAT 841  
DB 1005 CAGGAGGAGATTTGTGATCAGCGTGTCTCATGAGAGAACACCTCTCATTTGATTAAT 1064  
QY 842 GTCCTTTGGTGCAGCTGAGATCAGACCTTTTAAATGAAATAAATTGGAATAATGTACAACAG 901  
DB 1065 GTCCTTTGGTGCAGCTGAAATGAGACCTTTTAAATGAAATAAATTGGAATAATGTACAACAG 1124  
QY 902 CCTATGCTCTCTGTTATCCCAATGCAAGGCTTCCCAACACCTTTGGTGAATGATGAAA 961  
DB 1125 CCTATGCTCTCTGTTATCCCAATGCAAGGCTTCCCAACACCTTTGGTGAATGATGAAA 1184  
QY 962 CGCCTTCTATGATGGCCAGCACCTAAAGATTTTGTCTATGATGGCTTGGTCAATATAG 1021  
DB 1185 CGCCTTCTATGATGGCCAGCACCTAAAGATTTTGTCTATGATGGCTTGGTCAATATAG 1244  
QY 1022 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1081  
DB 1245 TTGAGGATGCTGTGGGTCAACACCAAGATCATATCAGGAGAAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAAGCCTAGAGTTCCACCTGCTCTTTTGAAGGACATATGTTTACTGTCTGGTC 1141

1305 ATTGTAAGCCTAGAGTTCCACCTGCGCCTCTTTGAAGGACATATGTTACTGTCGGTC 1364  
1142 TAGAGCCCTTCAGGATGGACCGGTACACCAACTTTGTAACTTGGAGAGCGCTGTAATG 1201  
1365 TAGAGCCCTTCAGGATGGACCGGTACACCAACTTTGTAACTTGGAGAGCGCTGTAATG 1424  
1202 TTGCAGATCAGGAAGTTTGTCTAACTCATCATGCGAGAACTATGAAGACCTTGT 1261  
1425 TTGCAGATCAGGAAGTTTGTCTAACTCATCATGCGAGAACTATGAAGACCTTGT 1484  
1262 GTCTTCCAAAGTCAGGTGGAATGGAGCCCGACAGTGTGGATGTCACATGAGATG 1321  
1485 GTCTTCCAAAGTCAGGTGGAATGGAGCCCGACAGTGTGGATGTCACATGAGATG 1544  
1322 GCATGCTAGATGTCAGGTGGAATGGAGCCCGACAGTGTGGATGTCACATGAGATG 1381  
1545 GCATGCTAGATGTCAGGTGGAATGGAGCCCGACAGTGTGGATGTCACATGAGATG 1604  
1382 ACATGCCAAAGTACCTTTGTGATGCACTCTCCCAATTTTGTCTCGAGCCAG 1441  
1605 ACATGCCAAAGTACCTTTGTGATGCACTCTCCCAATTTTGTCTCGAGCCAG 1664  
1442 TAAAGTGTGCCAAGGGAAGTGCATTTGCAATAGCATTTAGTCTGAGGAGGAGGAGCG 1501  
1665 TAAAGTGTGCCAAGGGAAGTGCATTTGCAATAGCATTTAGTCTGAGGAGGAGGAGCG 1724  
1502 ACTTCTTGGAGAGGCGCAGGAAGTAAAGAGTATGAGCTGTATGTTGGTTCATGGCTT 1561  
1725 ACTTCTTGGAGAGGCGCAGGAAGTAAAGAGTATGAGCTGTATGTTGGTTCATGGCTT 1784  
1562 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATAATGAGTGTGCACCCGGGCT 1621  
1785 TTGATGAAGAGGAGCAGGCAACAGAAACAGACACAAATAATGAGTGTGCACCCGGGCT 1844  
1622 ACCATCTGCTGTGAAGAACTGGGCTTTAATCCAAATGACATTTTGTGACCTTAATA 1681  
1845 ACCATCTGCTGTGAAGAACTGGGCTTTAATCCAAATGACATTTTGTGACCTTAATA 1904  
1682 TCCTAACCATTTGGGATGGGAATGGAGGACACAACTTTGTATGCCATTAATTTATCCATG 1741  
1905 TCCTAACCATTTGGGATGGGAATGGAGGACACAACTTTGTATGCCATTAATTTATCCATG 1964  
1742 CAACAAAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGAAGTCTTTTCCAACT 1801  
1965 CAACAAAGTCATTAAGAAACATTAACCTGGAGCCAGAAATAGTGAAGTCTTTTCCAACT 2024  
1802 TGTCTCTTCTCCGAGGAATGGAGCCATTCGAGAGCAATGCAATGGGTTTCTCTTT 1861  
2025 TGTCTCTTCTCCGAGGAATGGAGCCATTCGAGAGCAATGCAATGGGTTTCTCTTT 2084  
1862 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCCCTGTGT 1921  
2085 ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAACCTCCCTGTGT 2144  
1922 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACC 1981  
2145 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGACC 2204  
1982 CTGAGGCCATGAGAGCTCTTACGTTATGCTCCAGATCTCAAGGCAAGAGGGAAGAAAG 2041  
2205 CTGAGGCCATGAGAGCTCTTACGTTATGCTCCAGATCTCAAGGCAAGAGGGAAGAAAG 2264  
2042 TCATTTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACCTTTGAGTATGCCCTTG 2101  
2265 TCATTTCAGACTGATGAGTGGAGAAATGGCCCTGTGCAAGAACCTTTGAGTATGCCCTTG 2324  
2102 TGAAGGGCATTTGAAACATATTAATTTGAGGATCTGAGGAGCCAGGTTTAAACCAAAAA 2161  
2325 TGAAGGGCATTTGAAACATATTAATTTGAGGATCTGAGGAGCCAGGTTTAAACCAAAAA 2384  
2162 AATAATCCCGAATCTCTCAATATTAATTTGAAGGACCCCTGATGATGAATGAATAATTTG 2221

2385 AATATCCCGACCTCTCAATATATAATTTGAAGGACCCCTGATGATGAATGAATAATTTG 2444  
2222 GTGATCTTTTGGAGCTGGAAAAATTTTCTACCTCAGGTTATAAAGTCAGCCCGGTTA 2281  
2445 GTGATCTTTTGGAGCTGGAAAAATTTTCTACCTCAGGTTATAAAGTCAGCCCGGTTA 2504  
2282 TGAAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAAACAGAG 2341  
2505 TGAAGAGGCTGTTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAGAAACAGAG 2564  
2342 TGTCTTAAAGGCGACAGTAGAAGAGAGGAGCCCTTACAGGGCACCATCTGCTGGCCACTG 2401  
2565 TGTCTTAAAGGCGACAGTAGAAGAGAGGAGCCCTTACAGGGCACCATCTGCTGGCCACTG 2624  
2402 TTAAGAGGCGACCTGCAAG 2461  
2625 TTAAGAGGCGACCTGCAAG 2684  
2462 ATTTCCGAGTTATGATTTAGGAGTCATGACTCCATGATGATGAATGATGATGATGATG 2521  
2685 ATTTCCGAGTTATGATTTAGGAGTCATGACTCCATGATGATGAATGATGATGATGATG 2744  
2522 TTGACCAACAAAGCAGATATAATTTGGCCTGTGAGGACTCATCACTCTCTCCCTGGATGAA 2581  
2745 TTGACCAACAAAGCAGATATAATTTGGCCTGTGAGGACTCATCACTCTCTCCCTGGATGAA 2804  
2582 TGAATTTTGTCCCAAGGAAATGAGAGATTAAGGATTCATTTGTTGATTTGGAG 2641  
2805 TGAATTTTGTCCCAAGGAAATGAGAGATTAAGGATTCATTTGTTGATTTGGAG 2864  
2642 GAGCAACCACTTTCAAAAAACCAACAGAGAGTTAAATAGCTCCGAGATACAGTGCACCTG 2701  
2865 GAGCAACCACTTTCAAAAAACCAACAGAGAGTTAAATAGCTCCGAGATACAGTGCACCTG 2924  
2702 TAATCATGCTCTGAGAGGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAA 2761  
2925 TAATCATGCTCTGAGAGGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAA 2984  
2762 ATCTAAGAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC 2821  
2985 ATCTAAGAGATGAATATCTTTGAGGAAATCATGGAAGAAATATGAAGATATTAGACAGACC 3044  
2822 ATTATGATCTCTCAAGAGAGGAGATTAACCTTTAAGTCAAGCCAGAAAGTGGTT 2881  
3045 ATTATGATCTCTCAAGAGAGGAGATTAACCTTTAAGTCAAGCCAGAAAGTGGTT 3104  
2882 TCCAAATGATTTGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTGGAGCCAGG 2941  
3105 TCCAAATGATTTGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTGGAGCCAGG 3164  
2942 TCTTTGAAGACTATGACCTGACAGAGCTGGTGAAGTACATTTGACTGGAAGCCTTTCTTTG 3001  
3165 TCTTTGAAGACTATGACCTGACAGAGCTGGTGAAGTACATTTGACTGGAAGCCTTTCTTTG 3224  
3002 ATGTCCTGGAGCTCCGGGGCAAGTACCAGATTCGAGGCTTCCCAAGATATTATTAAGACA 3061  
3225 ATGTCCTGGAGCTCCGGGGCAAGTACCAGATTCGAGGCTTCCCAAGATATTATTAAGACA 3284  
3062 AATCAGATAGTGGAGAGGCGCAGAGAGGTTACGATGATGCCCAATATGCTGGAACAC 3121  
3285 AATCAGATAGTGGAGAGGCGCAGAGAGGTTACGATGATGCCCAATATGCTGGAACAC 3344  
3122 TGATTTAGTCAAAAAGAACTCCCGGGCCCGGGTGTGGTGGTGTCTGGCCAGCAGACAGAT 3181  
3345 TGATTTAGTCAAAAAGAACTCCCGGGCCCGGGTGTGGTGGTGTCTGGCCAGCAGACAGAT 3404  
3182 TCACAGAGCAGATTCACCTGTACGAGAGGCTGTGTGCTCCAGGCTGTCAGAGCCCATAG 3241  
3405 TCCAGAGCAGATTCACCTGTACGAGAGGCTGTGTGCTCCAGGCTGTCAGAGCCCATAG 3464  
3242 CCACCTTTCTATGGGTTTAAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3301  
3465 CCACCTTTCTATGGGTTTAAAGGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3524

QY 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACCTGGGCTCT 3361  
DB 3525 ACTGCTCTCAGACTTCATCGCTCCCTTGCAATCTGGCATCCGTGACTACCTGGGCTCT 3584  
QY 3362 TTGCCGTGCTCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3421  
DB 3585 TTGCCGTGCTCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3644  
QY 3422 ACTACAGCAGCATCATGGTCAAGCGCTGCGGACCGGCTGGGAGAGGCTTTGCAAGAG 3481  
DB 3645 ACTACAGCAGCATCATGGTCAAGCGCTGCGGACCGGCTGGGAGAGGCTTTGCAAGAG 3704  
QY 3482 AGCTCCATGAAGAGTTCCGCGAAGTGTGGGCTACTGTGGCAGTGAGCAGCTGGAG 3541  
DB 3705 AGCTCCATGAAGAGTTCCGCGAAGTGTGGGCTACTGTGGCAGTGAGCAGCTGGAG 3764  
QY 3542 TCGCAGACCTGCGAAGTTGCGGTACAAGGCGCATCCGCGGCTCCCTGGCTACCCGAGCC 3601  
DB 3765 TCGCAGACCTGCGAAGTTGCGGTACAAGGCGCATCCGCGGCTCCCTGGCTACCCGAGCC 3824  
QY 3602 AGCCCGACCAACCGAAGAGCTCAACATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3661  
DB 3825 AGCCCGACCAACCGAAGAGCTCAACATGTGGAGACTCGCAGACATCGAGCAGCTACAG 3884  
QY 3662 GCATTAGGTTAAAGATCAATGATCAATGCAATGGCAGCTGCTCAGGCTCTACT 3721  
DB 3885 GCATTAGGTTAAAGATCAATGATCAATGCAATGGCAGCTGCTCAGGCTCTACT 3944  
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTGCTGTGGGAGATTTCCAAAGGATCAGGTTGAGG 3781  
DB 3945 TCTCCAAATTTGAAGTCCAAATATTTGCTGTGGGAGATTTCCAAAGGATCAGGTTGAGG 4004  
QY 3782 ATTATGCAATTGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 3841  
DB 4005 ATTATGCAATTGAGGAAGACATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 4064  
QY 3842 TGGGATATGATACAGACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3875  
DB 4065 TGGGATATGATACAGACTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4098

## RESULT 4

US-09-347-878-4

; Sequence 4, Application US/09347878C

; Patent No. 6376210

; GENERAL INFORMATION:

; APPLICANT: Yuan, Chong

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES

; FILE REFERENCE: 25885-1651

; CURRENT APPLICATION NUMBER: US/09/347,878C

; CURRENT FILING DATE: 1999-07-06

; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 7122

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (287)..(4084)

; OTHER INFORMATION: Human methionine synthase cDNA

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: U75743/GenBank

US-09-347-878-4

Query Match 89.7%; Score 3517; DB 3; Length 7122;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

US-09-347-878-4

QY 2 GTCACTGTGAGAGCAGCTCTTCTGCGCGGCTCTGCGCAAGGAGAGACTCGACA 61

DB 225 GTCACTGTGAGAGCAGCTCTTCTGCGCGGCTCTGCGCAAGGAGAGACTCGACA 284

QY 62 ACATGTCACCCGCTCCAGACCTGTGCAACCCGAAAGGTCTGAAGAAAAACCTTCGGGG 121  
DB 285 ACATGTCACCCGCTCCAGACCTGTGCAACCCGAAAGGTCTGAAGAAAAACCTTCGGGG 344  
QY 122 ATGAGATCAATGCCATTCTGCAAGAGAGATTAATGGTCTGGATGGAGGATGGGAGCA 181  
DB 345 ATGAGATCAATGCCATTCTGCAAGAGAGATTAATGGTCTGGATGGAGGATGGGAGCA 404  
QY 182 TGATCAGCGGAGAGGCTTAACGAAAGACACTTCGAGGTTCAGGAATTTAAGATCATG 241  
DB 405 TGATCAGCGGAGAGGCTTAACGAAAGACACTTCGAGGTTCAGGAATTTAAGATCATG 464  
QY 242 CCAGCGCGCTGAAAGGCAACATGATGATTTAAGTATTAATCACTCAGCTGATGTCATTACC 301  
DB 465 CCAGCGCGCTGAAAGGCAACATGATGATTTAAGTATTAATCACTCAGCTGATGTCATTACC 524  
QY 302 AAATCCATAAGGAATATCTTGGCTGGGCGAGATATCAATGAAACAAATATCTTTAGCA 361  
DB 525 AAATCCATAAGGAATATCTTGGCTGGGCGAGATATCAATGAAACAAATATCTTTAGCA 584  
QY 362 GCACTAGTATTCGCCAAGCTGACTATGGCTTGAACACTTGGCCTACCGGATGAAACATGT 421  
DB 585 GCACTAGTATTCGCCAAGCTGACTATGGCTTGAACACTTGGCCTACCGGATGAAACATGT 644  
QY 422 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTTAAGA 481  
DB 645 GCTCTGCAGGAGTGGCCAGAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTTAAGA 704  
QY 482 GGTGTTGTGCAGGGGCTCTGGGTCGACATAAAGACACTCTCTGTGTCCCATCTGTGG 541  
DB 705 GGTGTTGTGCAGGGGCTCTGGGTCGACATAAAGACACTCTCTGTGTCCCATCTGTGG 764  
QY 542 AAAGCCGGAATTAAGGAACATCAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGS 601  
DB 765 AAAGCCGGAATTAAGGAACATCAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGS 824  
QY 602 CCAGAGGACTCTGGAATGGGGGTTGATATCTTACTCAATGAAACTATTTTGTGATCTG 661  
DB 825 CCAGAGGACTCTGGAATGGGGGTTGATATCTTACTCAATGAAACTATTTTGTGATCTG 884  
QY 662 CCAATGCCAAGCAGCCTTTGTTGCACTCCAAAATCTTTTGAAGGAATAATGTCCTCCC 721  
DB 885 CCAATGCCAAGCAGCCTTTGTTGCACTCCAAAATCTTTTGAAGGAATAATGTCCTCCC 944  
QY 722 GGCCTATCTTTATTTGAGGACGATGTTGATAAAGTGGGGGACTCTTTCGGGACAGA 781  
DB 945 GGCCTATCTTTATTTGAGGACGATGTTGATAAAGTGGGGGACTCTTTCGGGACAGA 1004  
QY 782 CAGGAGAGGATTTGTGATCAGCGTCTCTCATGAGAACCACTCTGCATTTGGATTTAAAT 841  
DB 1005 CAGGAGAGGATTTGTGATCAGCGTCTCTCATGAGAACCACTCTGCATTTGGATTTAAAT 1064  
QY 842 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATGTAACAG 901  
DB 1065 GTGCTTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATAATTTGAAATAATGTAACAG 1124  
QY 902 CCTATGTCCTCTGTTATCCCAATGCAAGGCTTCCCAACACCTTTGGTGATGATGAGAA 961  
DB 1125 CCTATGTCCTCTGTTATCCCAATGCAAGGCTTCCCAACACCTTTGGTGATGATGAGAA 1184  
QY 962 CGCCTTCTATGATGGCCCAAGCCTTAAGGATTTTGTATGATGCTGCTCAATATAG 1021  
DB 1185 CGCCTTCTATGATGGCCCAAGCCTTAAGGATTTTGTATGATGCTGCTCAATATAG 1244  
QY 1022 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1081  
DB 1245 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGAAATTTGCTGAAGCTGTGAAA 1304  
QY 1082 ATTGTAAGCCTAGAGTTCCACCTGGCACACTGCTTTTGAAGGACATATGTTACTGCTGTC 1141  
DB 1305 ATTGTAAGCCTAGAGTTCCACCTGGCACACTGCTTTTGAAGGACATATGTTACTGCTGTC 1364

QY	1142	TAGAGCCCTTCAGGATGGACCGGTACACCAACTTTGTATACATTTGGAGGCGCTGTAAATG	1201		2445	GTGATCTTTTGGAGCTGGAAAAATGTTTCTTACCTCAGGTATAAAGTCAGCCCGGTTA	2504		
Db	1365	TAGAGCCCTTCAGGATGGACCGGTACACCAACTTTGTATACATTTGGAGGCGCTGTAAATG	1424		QY	2282	TGAAGAAGGCTGTGTGGCCACCTTATCCCTTTTATGAAAAAGAAAGAGAAACCCAGAG	2341	
QY	1202	TTGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGCAAACTATGAAGAAGCCTTGT	1361		Db	2505	TGAAGNAGGCTGTGTGGCCACCTTATCCCTTTTATGAAAAAGAAAGAGAAACCCAGAG	2564	
Db	1425	TTGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGCAAACTATGAAGAAGCCTTGT	1484		QY	2342	TGCTTAAACGGCAGTAGAAGAGAGGACCTTTTACAGGGCCACCATGCTGTGGCCACTG	2401	
QY	1262	GTGTTCGCAAGTGCAGGTGGAAATGGAGCCAGGCTTGTGGATGCAACATGGATGATG	1321		Db	2565	TGCTTAAACGGCAGTAGAAGAGAGGACCTTTTACAGGGCCACCATGCTGTGGCCACTG	2624	
Db	1485	GTGTTCGCAAGTGCAGGTGGAAATGGAGCCAGGCTTGTGGATGCAACATGGATGATG	1544		QY	2402	TTAAAGCGCAGCTGCACGACATAGGCAAGAAACATAGTTCGAGTAGTCCCTTGGCTGCAATA	2461	
QY	1322	GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTGCATTAATTTTCCGAGCCAG	1381		Db	2625	TTAAAGCGCAGCTGCACGACATAGGCAAGAAACATAGTTCGAGTAGTCCCTTGGCTGCAATA	2684	
Db	1545	GCATGCTAGATGCTCCAAAGTGCATGACCAAGATTTTGCATTAATTTTCCGAGCCAG	1604		QY	2462	ATTTCGAGTTTATGATTTAGGAGTCATGACTCCATGTGATAGATATCTGAAAGCTGCTC	2521	
QY	1382	ACATCGAAAGTACCTTTTGTGATCGACTCCCTCAATTTTGTGATTAAGAGCTGGGT	1441		Db	2685	ATTTCGAGTTTATGATTTAGGAGTCATGACTCCATGTGATAGATATCTGAAAGCTGCTC	2744	
Db	1605	ACATCGAAAGTACCTTTTGTGATCGACTCCCTCAATTTTGTGATTAAGAGCTGGGT	1664		QY	2522	TTGACCAAAAGCAGATATATTTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2581	
QY	1442	TAAAGTGTCTCCAAAGGAAGTGCATTTCAATAGCAATTAAGTGCATTAAGAGGAGGAGACG	1501		Db	2745	TTGACCAAAAGCAGATATATTTGGCTGTGAGGACTCATCACTCTCTCCCTGGATGAAA	2804	
Db	1665	TAAAGTGTCTCCAAAGGAAGTGCATTTCAATAGCAATTAAGTGCATTAAGAGGAGGAGACG	1724		QY	2582	TGATTTTTCGTCGCAAGGAATGGAGAGATTAAGTATTAAGGATTCCTATTTGATTTGGAG	2641	
QY	1502	ACTTCTTGGAGAGGCCAGGAAGTTTAAAGATGATGAGGCTCTATGCTGTCTATGGCTT	1561		Db	2805	TGATTTTTCGTCGCAAGGAATGGAGAGATTAAGTATTAAGGATTCCTATTTGATTTGGAG	2864	
Db	1725	ACTTCTTGGAGAGGCCAGGAAGTTTAAAGATGATGAGGCTCTATGCTGTCTATGGCTT	1784		QY	2642	GAGCAACCACTTCAAAAACCCACACAGCAGTGTAAAAATAGCTCCGAGATACAGTGCACTTG	2701	
QY	1562	TTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCT	1621		Db	2865	GAGCAACCACTTCAAAAACCCACACAGCAGTGTAAAAATAGCTCCGAGATACAGTGCACTTG	2924	
Db	1785	TTGATGAAGAGGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCT	1844		QY	2702	TAATCCATGCTCTGGAAGCGCTCAAGAGTGTGGTGTGTGTTTCCAGCTGTTTAGATGAAA	2761	
QY	1622	ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCTTAATA	1681		Db	2925	TAATCCATGCTCTGGAAGCGCTCAAGAGTGTGGTGTGTGTTTCCAGCTGTTTAGATGAAA	2984	
Db	1845	ACCATCTGCTTGTGAAAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCTTAATA	1904		QY	2762	ATCTAAAGGATGAATACTTTTGGAGAAATCATGGAAGAAATATGAAGATTTAGACAGACC	2821	
QY	1682	TCTAACCATTTGGATGGAGGAAACAACTCTGTATGATGCAATTAATTTTATCCATG	1741		Db	2985	ATCTAAAGGATGAATACTTTTGGAGAAATCATGGAAGAAATATGAAGATTTAGACAGACC	3044	
Db	1905	TCTAACCATTTGGATGGAGGAAACAACTCTGTATGATGCAATTAATTTTATCCATG	1964		QY	2822	ATTATGAGTCTCTCAAGAGAGGAGATACTTACCTTAACTCAAGCCAGCAAAAAGTGGTT	2881	
QY	1742	CAACAAAGTCTTAAAGAAACATTAACCTGAGCCAGAAATAGTGGAGGCTTTTCCAACT	1801		Db	3045	ATTATGAGTCTCTCAAGAGAGGAGATACTTACCTTAACTCAAGCCAGCAAAAAGTGGTT	3104	
Db	1965	CAACAAAGTCTTAAAGAAACATTAACCTGAGCCAGAAATAGTGGAGGCTTTTCCAACT	2024		QY	2882	TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGAGCCAGG	2941	
QY	1802	TGTCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGCTTTTCTTTT	1861		Db	3105	TCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTATTTGGAGCCAGG	3164	
Db	2025	TGTCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAGCAATGCAATGGGCTTTTCTTTT	2084		QY	2942	TCTTTGAAGACTATGACCTGCAAGAGCTGGTGGACTACATTTGACTGGAGCCCTTTCTTTG	3001	
QY	1862	ACCATGCAATCAAGTCTGGCATGGACATGGAGATAGTGAATGCTGGAAACCTCCCTGTGT	1921		Db	3165	TCTTTGAAGACTATGACCTGCAAGAGCTGGTGGACTACATTTGACTGGAGCCCTTTCTTTG	3224	
Db	2085	ACCATGCAATCAAGTCTGGCATGGACATGGGGATAGTGAATGCTGGAAACCTCCCTGTGT	2144		QY	3002	ATGCTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGCACA	3061	
QY	1922	ATGATGATATCCATAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGCC	1981		Db	3225	ATGCTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGGCTTCCCAAGATATTTAAGCACA	3284	
Db	2145	ATGATGATATCCATAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGCC	2204		QY	3062	AAACAGTAGTGGAGAGGCGCAGGAAGGCTCTACGATGATGCCCAATATGCTGAAACAC	3121	
QY	1982	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCCACAGAGGGAAGAAAG	2041		Db	3285	AAACAGTAGTGGAGAGGCGCAGGAAGGCTCTACGATGATGCCCAATATGCTGAAACAC	3344	
Db	2205	CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCCACAGAGGGAAGAAAG	2264		QY	3122	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGGTGTGGTGGTCTGGCCAGCACAGAGTA	3181	
QY	2042	TCATTACAGATGATGAGGAAATGGCCCTGTGGAAGAACCCCTTGTAGTATGCCCTTG	2101		Db	3345	TGATTTAGTCAAAAAGAACTCCGGGCCCGGGGTGTGGTGGTCTGGCCAGCACAGAGTA	3404	
Db	2265	TCATTACAGATGATGAGGAAATGGCCCTGTGGAAGAACCCCTTGTAGTATGCCCTTG	2324		QY	3182	TCCAGAGCCACATTCACCTGTACGAGAGGCTGTGTGCCCGCAGGCTGAGAGGCCCATAG	3241	
QY	2102	TGAAGGCCATGAAACATATTTTGAAGATCTGAGGAGCCAGGTTAAACCAAAAAA	2161		Db	3405	TCCAGAGCCACATTCACCTGTACGAGAGGCTGTGTGTGCCCGCAGGCTGAGAGGCCCATAG	3464	
Db	2325	TGAAGGCCATGAAACATATTTTGAAGATCTGAGGAGCCAGGTTAAACCAAAAAA	2384		QY	3242	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATCT	3301	
QY	2162	AATATCCCGACCTCTCAATATATTAAGAGGACCCCTGTATGAATGAATGAATTTGTTG	2221		Db	3465	CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAGGACTCTGCCAGCAGGAGCCATCT	3524	
Db	2385	AATATCCCGACCTCTCAATATATTAAGAGGACCCCTGTATGAATGAATGAATTTGTTG	2444		QY	3302	ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCAATCCGTGACTACCTGGGCTGT	3361	
QY	2222	GTGATCTTTTGGAGCTGGAAAAATGTTTCTACCTCAGGTTATAAAGTCAGCCCGGGTTA	2281						



Db 3525 ACTGCCTCTCAGACTTCATCGCTCCCTTGGCACTTCTGGCACTCCGTGACTACCTGGGCTGT 3584  
Qy 3362 TTCCCGTTGCTCTGCTTTGGGGTAAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3421  
Db 3585 TTCCCGTTGCTCTGCTTTGGGGTAAAGAGCTGAGCAAGGCTATGAGGATGATGGTGACG 3644  
Qy 3422 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAAGAG 3481  
Db 3645 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAAGAG 3704  
Qy 3482 AGCTCATGAAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3541  
Db 3705 AGCTCATGAAAGAGTTTCGCCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGGACG 3764  
Qy 3542 TCGCAGACCTGCGAAGGTTGGGTACAAGGGCAATCCGCCCGGCTCTGGCTACCCAGCC 3601  
Db 3765 TCGCAGACCTGCGCAGGCTCGGTACAAAGGGCAATCCGCCCGGCTCTGGCTACCCAGCC 3824  
Qy 3602 AGCCGACACACCGGAGGCTCACCAGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3661  
Db 3825 AGCCGACACACCGGAGGCTCACCAGTGGAGACTCGCAGACATCGAGCAGCTCTACAG 3884  
Qy 3662 GCATTAGTTAAACAGAACTATTAGCAATGGCAGCTCTTTCAGCAGTCTCAGGCTCTACT 3721  
Db 3885 GCATTAGTTAAACAGAACTATTAGCAATGGCAGCTCTTTCAGCAGTCTCAGGCTCTACT 3944  
Qy 3722 TCTCCAAATTTGAAGTCAAAATATTTTGTCTGGGGAAAGATTTCCAGGATCAGGTTGAGG 3781  
Db 3945 TCTCCAAATTTGAAGTCAAAATATTTTGTCTGGGGAAAGATTTCCAGGATCAGGTTGAGG 4004  
Qy 3782 ATTATGCAATTCAGGAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCAATTT 3841  
Db 4005 ATTATGCAATTCAGGAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCAATTT 4064

## RESULT 5

US-09-577-266-2  
; Sequence 2, Application US/09577266  
; Patent No. 6912492  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, William G.  
; APPLICANT: Stenroos, Edward S.  
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS  
; FILE REFERENCE: 601-1-057N  
; CURRENT APPLICATION NUMBER: US/09/577,266  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/136,198  
; PRIOR FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 7122  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-577-266-2

Query Match 89.7%; Score 3517; DB 3; Length 7122;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3867; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 GTCACCTGTGAGAGCAGCTCTCTGCGCGCCCTCTGGCAAGGAGGAGACTCGACA 61  
Db 225 GTCACCTGTGAGAGCAGCTCTCTGCGCGCCCTCTGGCAAGGAGGAGACTCGACA 284  
Qy 62 ACATGTCAACCGCGCTCAAGACCTGTGCGAACCCGAGGCTGGAAGAAACCCCTGCGGG 121  
Db 285 ACATGTCAACCGCGCTCAAGACCTGTGCGAACCCGAGGCTGGAAGAAACCCCTGCGGG 344

Qy 122 ATGAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTGCTGATGAGGGATGGGACCA 181  
Db 345 ATGAGATCAATGCCATTTCTGCAGAGAGGATTTATGCTGCTGATGAGGGATGGGACCA 404  
Qy 182 TGATCAGCGGAGAGCTTAAACGAAGAACATTTCCGAGGTGAGGATTTAAAGATCATG 241  
Db 405 TGATCAGCGGAGAGCTTAAACGAAGAACATTTCCGAGGTGAGGATTTAAAGATCATG 464  
Qy 242 CGAGCGCTGAAAGCAACATGATCAATTTTAAGTATTAATCAGCTGATGCTATTACC 301  
Db 465 CGAGCGCTGAAAGCAACATGATCAATTTTAAGTATTAATCAGCTGATGCTATTACC 524  
Qy 302 AAATCCATAAGGAATACTTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTTAGCA 361  
Db 525 AAATCCATAAGGAATACTTGTGGCTGGGCGAGATATCAATGAAACAAATATCTTTTAGCA 584  
Qy 362 GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAAACATTTGGCCTACCGAGTGAACATGT 421  
Db 585 GCACTAGTATTGCCCAAGCTGACTATGGCTTTGAAACATTTGGCCTACCGAGTGAACATGT 644  
Qy 422 GCTCTCAGAGTGGCCAGAAAAGCTGCGGAGGATTAATCTCTCCAGACAGGAATTAAGA 481  
Db 645 GCTCTCAGAGTGGCCAGAAAAGCTGCGGAGGATTAATCTCTCCAGACAGGAATTAAGA 704  
Qy 482 GGTTCGTGCGAGGGCTCTGGGTCCGACTAAAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 705 GGTTCGTGCGAGGGCTCTGGGTCCGACTAAAGACACTCTCTGTGTCCCATCTGTGG 764  
Qy 542 AAAGCGCGGATATAGGAACATCACAATTTGATGAGCTTGTGGAAGCATACCAAGACAGG 601  
Db 765 AAAGCGCGGATATAGGAACATCACAATTTGATGAGCTTGTGGAAGCATACCAAGACAGG 824  
Qy 602 CMAAGGACTCTGGATGGCGGGTGTATCTTACTCATTTGAAACTATTTTGTGACTG 661  
Db 825 CMAAGGACTCTGGATGGCGGGTGTATCTTACTCATTTGAAACTATTTTGTGACTG 884  
Qy 662 CCAATGCCAAGGCGAGCTTGTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
Db 885 CCAATGCCAAGGCGAGCTTGTGTTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 944  
Qy 722 GGCCTATCTTTATTTTCAGGAGCATCGTTGATAAAAGTGGCGGACTCTTTCCGGAAGA 781  
Db 945 GGCCTATCTTTATTTTCAGGAGCATCGTTGATAAAAGTGGCGGACTCTTTCCGGAAGA 1004  
Qy 782 CAGGAGAGGATTTGTTCATCAGGCTGTCTCATGAGAACCACTCTGCTGATGGAATTAAT 841  
Db 1005 CAGGAGAGGATTTGTTCATCAGGCTGTCTCATGAGAACCACTCTCATTTGGATTAAT 1064  
Qy 842 GTGCTTTGGTGCAGCTGAGATGAGACCTTTTATGAAATAATTTGAAATAATGACACAG 901  
Db 1065 GTGCTTTGGTGCAGCTGAAATGAGACCTTTTATGAAATAATTTGAAATAATGACACAG 1124  
Qy 902 CCTATGTCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTGGTGAATGATGAAA 961  
Db 1125 CCTATGTCTCTGTTATCCCAATGCAAGTCTTCCCAACACCTTTGGTGAATGATGAAA 1184  
Qy 962 CGCCTTCTATGATGGCCAGACCTTAAAGATTTTGGTATGATGGCTGGTCAATATAG 1021  
Db 1185 CGCCTTCTATGATGGCCAGACCTTAAAGATTTTGGTATGATGGCTGGTCAATATAG 1244  
Qy 1022 TTGGAGGATGCTGTGGGTCAACACCATCATATCAGGGGAAATTTGCTGAAAGCTGTGAAA 1081  
Db 1245 TTGGAGGATGCTGTGGGTCAACACCATCATATCAGGGGAAATTTGCTGAAAGCTGTGAAA 1304  
Qy 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTCTCTTTTGAAGGACATATGTTACTGTCTGTC 1141  
Db 1305 ATTGTAAAGCTTAGAGTTCCACCTGCCACTCTCTTTTGAAGGACATATGTTACTGTCTGTC 1364  
Qy 1142 TAGAGCCCTTCAGGATTTGGAACCGTACACCACTTTGTTTAACTTGGAGAGCGCTGTAATG 1201  
Db 1365 TAGAGCCCTTCAGGATTTGGAACCGTACACCACTTTGTTTAACTTGGAGAGCGCTGTAATG 1424  
Qy 1202 TTGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGGAAACTATGAAAGAGCCTTGT 1261



1425 |||||TTGCGAGATCAAGGAGTTTGTCTAACTCATCTGGCAGGAACATATGAAGAAGCCTTGT 1484  
1262 GTGTTGCCAAAGTCAGGTGGAATGGAGGCCAGGTGTTGGATGTCAACATGATGATG 1321  
1485 GTGTTGCCAAAGTCAGGTGGAATGGAGGCCAGGTGTTGGATGTCAACATGATGATG 1544  
1322 GCATGCTAGATGTTCCAAAGTCAGATGACCAAGATTTCCAACTTAATTTGCTTCCAGCCAG 1381  
1545 GCATGCTAGATGTTCCAAAGTCAGATGACCAAGATTTCCAACTTAATTTGCTTCCAGCCAG 1604  
1382 ACATCGCAAGGTACCTTTGTGCACTGACCTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1441  
1605 ACATCGCAAGGTACCTTTGTGCACTGACCTCTCCAAATTTTGTGCTGATTTGAAGCTGGGT 1664  
1442 TAAAGTGTCTCCAAAGGAAGTGCATTTGTCAATAGCATTAGTCTGAAGGAAGGAGGACG 1501  
1665 TAAAGTGTCTCCAAAGGAAGTGCATTTGTCAATAGCATTAGTCTGAAGGAAGGAGGACG 1724  
1502 ACTTCTTGGAGAGGCCAGGAAGATTAAGAAAGTATAGGAGCTGCTATGGTGGTCAATGGCTTT 1561  
1725 ACTTCTTGGAGAGGCCAGGAAGATTAAGAAAGTATAGGAGCTGCTATGGTGGTCAATGGCTTT 1784  
1562 TTGATGAAGAGGACGACGCAACAGAAACAGACACAAATAACAGAGTGTGCACCCGGGSCCT 1621  
1785 TTGATGAAGAGGACGACGCAACAGAAACAGACACAAATAACAGAGTGTGCACCCGGGSCCT 1844  
1622 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGGACCTTAATA 1681  
1845 ACCATCTGCTTGTGAAAAAATCTGGGCTTTAATCCAAATGACATTTATTTTGAACCTTAATA 1904  
1682 TCCTAACATTTGGACATGGAATGGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 1741  
1905 TCCTAACATTTGGACATGGAATGGAGGAACACAACTTGTATGCCATTAATTTTATCCATG 1964  
1742 CAACAAAGTCATTAAAGAAACATTTACCTGGAGCCCAATTAAGTGGAGGTCTTTTCCAACT 1801  
1965 CAACAAAGTCATTAAAGAAACATTTACCTGGAGCCCAATTAAGTGGAGGTCTTTTCCAACT 2024  
1802 TGTCTCTTCTCCTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGCAATGGGCTTTTCCCTTT 1861  
2025 TGTCTCTTCTCCTCCGAGGAATGGAAGCCATTTGAGAAAGCAATGCAATGGGCTTTTCCCTTT 2084  
1862 ACCATGCAATCAAGTCTGGCAATGGACATGAGATGAGTGAATGCTGGAAACCTCCCTGTGT 1921  
2085 ACCATGCAATCAAGTCTGGCAATGGACATGAGTGAATGCTGGAAACCTCCCTGTGT 2144  
1922 ATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAGATCTCATCTGGAATTAAGACC 1981  
2145 ATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAGATCTCATCTGGAATTAAGACC 2204  
1982 CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2041  
2205 CTGAGGCCACTGAGAAGCTCTTACGTTATGCCAGACTCAAGGCACAGGAGGGAAGAAAG 2264  
2042 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG 2101  
2265 TCATTCAGACTGATGAGTGGAGAAATGGCCCTGTGGAAGAACGCTTTGAGTATGCCCTTG 2324  
2102 TGAAGGCCATTGAAAAACATATTTATGAGATACTGAGGAAGCCAGGTTAAACCAAAAAA 2161  
2325 TGAAGGCCATTGAAAAACATATTTATGAGATACTGAGGAAGCCAGGTTAAACCAAAAAA 2384  
2162 AATATCCCGACCTCTCAATATATTAAGGACCCCTGATGAATGGAATGGAATGGAATTTGTTG 2221  
2385 AATATCCCGACCTCTCAATATATTAAGGACCCCTGATGAATGGAATGGAATTTGTTG 2444  
2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATTAAGTCAAGCCCGGTTA 2281  
2445 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATTAAGTCAAGCCCGGTTA 2504  
2282 TGAAGAAGGCTGTTGGCCACTTATCCCTTTTCAATGGAAAAAGAAAGAAACCAAGAG 2341

2505 TGAAGAAGGCTGTTGGCCACCTTATCCCTTTTCAATGAAAAAGAAAGAAACCAAGAG 2564  
2342 TGCTTAAACGCAACAGTAGAAGAGAGGACCTTTTACAGGCAACCATCGTGTCTGCCACTG 2401  
2565 TGCTTAAACGCAACAGTAGAAGAGAGGACCTTTTACAGGCAACCATCGTGTCTGCCACTG 2624  
2402 TTAAGAGCGACGTCGACGACATAGGCAAGACATAGTTGGAGTAGTCTTGGTGGTCAATA 2461  
2625 TTAAGAGCGACGTCGACGACATAGGCAAGACATAGTTGGAGTAGTCTTGGTGGTCAATA 2684  
2462 ATTTCCGAGTTATTTAGTTTGGAGTCACTGACTCCATGTGATAAGATACTGAAAAAGCTGCTC 2521  
2685 ATTTCCGAGTTATTTAGTTTGGAGTCACTGACTCCATGTGATAAGATACTGAAAAAGCTGCTC 2744  
2522 TTGACCAAAAGCAGATATAAATGGCCTGTGAGGACTCACTCCTCTCCCTGGATGAAA 2581  
2745 TTGACCAAAAGCAGATATAAATGGCCTGTGAGGACTCACTCCTCTCCCTGGATGAAA 2804  
2582 TGAATTTTGTGTCGAAGGAATGGAGAGATAGCTATAGGATTTCCATTTGTTGATTGGAG 2641  
2805 TGAATTTTGTGTCGAAGGAATGGAGAGATAGCTATAGGATTTCCATTTGTTGATTGGAG 2864  
2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTG 2701  
2865 GAGCAACCACTTCAAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTG 2824  
2702 TAATCCTGCTCTGGACGCTCCAGAGTGTGGTGTGTTTCCAGCTGTTTATAGATGAAA 2761  
2925 TAATCCTGCTCTGGACGCTCCAGAGTGTGGTGTGTTTCCAGCTGTTTATAGATGAAA 2984  
2762 ATCTAAAGATGATTAATTTTGGAGGAATCATGGAAGAAATATGAAGATATTTAGACAGACC 2821  
2985 ATCTAAAGATGATTAATTTTGGAGGAATCATGGAAGAAATATGAAGATATTTAGACAGACC 3044  
2822 ATTTACAGTCTCTCAAGGAGAGGAGATACCTTACCCTTTAAGTCAAGCCAGAAAAAGTGGTT 2881  
3045 ATTTACAGTCTCTCAAGGAGAGGAGATACCTTACCCTTTAAGTCAAGCCAGAAAAAGTGGTT 3104  
2882 TCCAAAATGGAATGGCTGTCTGAACTCACCCAGTGAAGCCACGCTTTTATTTGGAGCCAGG 2941  
3105 TCCAAAATGGAATGGCTGTCTGAACTCACCCAGTGAAGCCACGCTTTTATTTGGAGCCAGG 3164  
2942 TCTTTGAAGACTATGACCTGCGAAGCTGGTGGACATACTTGAATGCTGGAAGCCCTTTCTTTG 3001  
3165 TCTTTGAAGACTATGACCTGCGAAGCTGGTGGACATACTTGAATGCTGGAAGCCCTTTCTTTG 3224  
3002 ATGCTCTGGCAGCTCCGGGCGAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTAAAGCA 3061  
3225 ATGCTCTGGCAGCTCCGGGCGAAGTACCCGAAATCGAGGCTTCCCAAGATATTTTAAAGCA 3284  
3062 AAACAGTAGTGGAGAGGCCAGGAAGGCTCTACGATGATGCCCAATATGCTGAAACACAC 3121  
3285 AAACAGTAGTGGAGAGGCCAGGAAGGCTCTACGATGATGCCCAATATGCTGAAACACAC 3344  
3122 TGAATTAAGTCAAAAGAAATCTCCGGGCCCCGGGTGTTGTTGGTCTTGGCCAGACACAGATTA 3181  
3345 TGAATTAAGTCAAAAGAAATCTCCGGGCCCCGGGTGTTGTTGGTCTTGGCCAGACACAGATTA 3404  
3182 TCCAGACGACATTCACCTGTACGCGAGAGGCTGCTGTGCCCGAGGCTGAGAGCCCATAG 3241  
3405 TCCAGACGACATTCACCTGTACGCGAGAGGCTGCTGTGCCCGAGGCTGAGAGCCCATAG 3464  
3242 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAAAGGACTCTGCGACAGCGAGGCCATACT 3301  
3465 CCACCTTCTATGGGTTAAGGCAACAGGCTGAGAAAGGACTCTGCCAGACGGAGCCATACT 3524  
3302 ACTGCTCTCAGACTTTCATCGCTCCCTTGGATTTCTGGCATCCGTGACTACCTGGGCTGT 3361  
3525 ACTGCTCTCAGACTTTCATCGCTCCCTTGGATTTCTGGCATCCGTGACTACCTGGGCTGT 3584  
3362 TTGCCGCTTCCCTGCTTTGGGTTAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3421  
3585 TTGCCGCTTCCCTGCTTTGGGTTAAGAGCTGAGCAAGGCTTATGAGGATGATGGTGACG 3644

QY 3422 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGCTGGCAGAGGCGCTTTGCGAAG 3481  
DB 3645 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGCTGGCAGAGGCGCTTTGCGAAG 3704  
QY 3482 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGCGCTACTGTGGCAGTGGACAGCTGGACG 3541  
DB 3705 AGCTCCATGAAGAGTTCCGCCGAGAACTGTGGCGCTACTGTGGCAGTGGACAGCTGGACG 3764  
QY 3542 TCGCAGACCTGCGAAGGTTCCGGTACAAAGGCAATCCGCCCGCTCTCTGGCTACCCGAGCC 3601  
DB 3765 TCGCAGACCTGCGCAGGCTCGGTTACAAAGGCAATCCGCCCGCTCTCTGGCTACCCGAGCC 3824  
QY 3602 AGCCGACACACCGAGAACTCACCATGTGGAGATCGCAGACATCGAGCACTCTACAG 3661  
DB 3825 AGCCGACACACCGAGAACTCACCATGTGGAGATCGCAGACATCGAGCACTCTACAG 3884  
QY 3662 GCATTAGTTAAAGAAATCAATAGCAATGGCAGCTCTCTCAGCAGTCTCAGGCTCTACT 3721  
DB 3885 GCATTAGTTAAAGAAATCAATAGCAATGGCAGCTCTCTCAGCAGTCTCAGGCTCTACT 3944  
QY 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTCTGTGGGAAAGATTTCCAAAGATCAGGTTGAGG 3781  
DB 3945 TCTCCAAATTTGAAGTCCAAATATTTTGTCTGTGGGAAAGATTTCCAAAGATCAGGTTGAGG 4004  
QY 3782 ATTATGCAATTTGAGGAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 3841  
DB 4005 ATTATGCAATTTGAGGAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTTGGACCCATTT 4064  
QY 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTT 3875  
DB 4065 TGGGATATGATACAGACTAACTTTTTTTTTTTT 4098

RESULT 6  
US-09-347-878-6  
; Sequence 6, Application US/09347878C  
; Patent No. 6376210  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Chong  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES  
; FILE REFERENCE: 25885-1651  
; CURRENT APPLICATION NUMBER: US/09/347,878C  
; CURRENT FILING DATE: 1999-07-06  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 7224  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (395)..(4192)  
; FEATURE:  
; OTHER INFORMATION: Human methionine synthase  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: U73338/GenBank  
US-09-347-878-6

Query Match 88.4%; Score 3466; DB 3; Length 7224;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3666; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GTCACTGTGTGAGAGCAGTCTTCTCTGCCCGCCCTCTCGCAAGGAGGAGACTCGACA 61  
DB 333 GTCACTGTGTGAGAGCAGTCTTCTCTGCCCGCCCTCTCGCAAGGAGGAGACTCGACA 392  
QY 62 ACATGTACCCGCGCTCCAAAGCTGTCTGCAACCCGGAAGGTTCTGAAGAAACCTCGGG 121  
DB 393 ACATGTACCCGCGCTCCAAAGCTGTCTGCAACCCGGAAGGTTCTGAAGAAACCTCGGG 452  
QY 122 ATGAGATCAATGCCATCTCTGAGAAAGAGGATTTATGGTCTCGATGGAGGATGGGACCA 181

DB 453 ATGAGATCAATGCCATCTTTCGCAAGAGAGGATTTATGGTCTCGATGGAGGATGGGACCA 512  
QY 182 TGAATCCAGCGGAGAGCTAAACGAGACACTTCGAGGTGAGGAATTTAAAGATCATG 241  
DB 513 TGAATCCAGCGGAGAGCTAAACGAGAACTTCGAGGTGAGGAATTTAAAGATCATG 572  
QY 242 CAGGCGCGCTGAAAGGCAACAAATGACATTTTAAAGTATAACTCAGCCTGATGCTATTACC 301  
DB 573 CAGGCGCGCTGAAAGGCAACAAATGACATTTTAAAGTATAACTCAGCCTGATGCTATTACC 632  
QY 302 AAATCCATTAAGGAATATCTTGGCTGGGCGAGATATCAATTGAAACAAATATCTTTAGCA 361  
DB 633 AAATCCATTAAGGAATATCTTGGCTGGGCGAGATATCAATTGAAACAAATATCTTTAGCA 692  
QY 362 GCACTAGTATTCGCCAAGCTGACTATGGCTTTGACACTTGGCTTACCCGATCGAATGATGT 421  
DB 693 GCACTAGTATTCGCCAAGCTGACTATGGCTTTGACACTTGGCTTACCCGATCGAATGATGT 752  
QY 422 GCTCTCAGAGAGTGGCCAGAAAAGCTGCGAGAGGTAACTCTCCAGACAGGAATTTAAGA 481  
DB 753 GCTCTCAGAGAGTGGCCAGAAAAGCTGCGAGAGGTAACTCTCCAGACAGGAATTTAAGA 812  
QY 482 GGTTCGTGCAAGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCACTGTGTGG 541  
DB 813 GGTTCGTGCAAGGCTCTGGGTCCGACTAAATAAGACACTCTCTGTGTCCCACTGTGTGG 872  
QY 542 AAAGGCGGATTAAGBAACATCACAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 601  
DB 873 AAAGGCGGATTAAGBAACATCACAATTTGATGAGCTTGTGAAGCATACCAAGAGCAGG 932  
QY 602 CCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGACTG 661  
DB 933 CCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTTGAAACTATTTTGTGACTG 992  
QY 662 CCAATGCCAAGGCGAGCTTTTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721  
DB 993 CCAATGCCAAGGCGAGCTTTTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052  
QY 722 GGCCTATCTTTATTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCGGGACAGA 781  
DB 1053 GGCCTATCTTTATTCAGGACGATCGTTGATAAAGTGGCGGACTCTTTCGGGACAGA 1112  
QY 782 CAGGAGAGGATTTGTTCATCAGCGTGTCTCATGAGAACCACTCTCGATTTGATTAAT 841  
DB 1113 CAGGAGAGGATTTGTTCATCAGCGTGTCTCATGAGAACCACTCTCGATTTGATTAAT 1172  
QY 842 GTGCTTTGGGTGCGAGTGAAGTACGACCTTTTATGAAATPAATTTGAAATTTGAAACAG 901  
DB 1173 GTGCTTTGGGTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1232  
QY 902 CCTATGTCCTCTGTTATCCCAATGCGAGTCTTCCCAACACCTTTGGTGAATGATGAAA 961  
DB 1233 CCTATGTCCTCTGTTATCCCAATGCGAGTCTTCCCAACACCTTTGGTGAATGATGAAA 1292  
QY 962 CGCTTCTATGATGGCCCAAGCACTAAAGGATTTTGTCTATGGATGGTCTGATCAATATAG 1021  
DB 1293 CGCTTCTATGATGGCCCAAGCACTAAAGGATTTTGTCTATGGATGGTCTGATCAATATAG 1352  
QY 1022 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGGAAATTTGCTGAAGCTGTGAAA 1081  
DB 1353 TTGAGGATGCTGTGGGTCAACACAGATCATATCAGGGGAAATTTGCTGAAGCTGTGAAA 1412  
QY 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGTC 1141  
DB 1413 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGCTGTC 1472  
QY 1142 TAGAGCCCTTCAGGATTTGGACCGGTACACCAACTTTTGTAACTTTGAGAGCGCTGTAATG 1201  
DB 1473 TAGAGCCCTTCAGGATTTGGACCGGTACACCAACTTTTGTAACTTTGAGAGCGCTGTAATG 1532  
QY 1202 TTGAGGATCAAGGAAGTTTGTAAACTCATATGCGAGGAAACTATGAGAAAGCTTTGT 1261  
DB 1533 TTGAGGATCAAGGAAGTTTGTAAACTCATATGCGAGGAAACTATGAGAAAGCTTTGT 1592

1262 GTGTTGCCAAAGTGCAGGTGGAAATGGAGAGCCAGGTGTTGGATGTCAACATCGATGATG 1321  
1593 GTGTTGCCAAAGTGCAGGTGGAAATGGAGAGCCAGGTGTTGGATGTCAACATCGATGATG 1652  
1322 GCATGCTAGATGTTCCAAAGTGCATGACACAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1381  
1653 GCATGCTAGATGTTCCAAAGTGCATGACACAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1712  
1382 ACATCGCAAGGTTACCTTTGTGTCATCGACTCTCCCAATTTTGTGCTGATGTAAGCTGGGT 1441  
1713 ACATCGCAAGGTTACCTTTGTGTCATCGACTCTCCCAATTTTGTGCTGATGTAAGCTGGGT 1772  
1442 TAAAGTGTCTCCAAAGGAGTGCATTTGTCAATAGCAATTAGTCTGAAGGAAGGAGAGAG 1501  
1773 TAAAGTGTCTCCAAAGGAGTGCATTTGTCAATAGCAATTAGTCTGAAGGAAGGAGAGAG 1832  
1502 ACTTCTTGGAGAGAGGCGCAGGAAGATTAAAGATGATGAGCTGTATGTTGGTGTCAATGCTT 1561  
1833 ACTTCTTGGAGAGGCGCAGGAAGATTAAAGATGATGAGCTGTATGTTGGTGTCAATGCTT 1892  
1562 TTGATGAAGAGGACGAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCT 1621  
1893 TTGATGAAGAGGACGAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCACCCGGGCGCT 1952  
1622 ACCATCTGCTTGTGAAGAACTGGGCTTTAATCCAAATGACATTTATTTTGGACCTTAATA 1681  
1953 ACCATCTGCTTGTGAAGAACTGGGCTTTAATCCAAATGACATTTATTTTGGACCTTAATA 2012  
1682 TCCCTAACCATTTGGGACTGGAAATGGAGGACACAACTTTGTATGCCATTTATTTATCCATG 1741  
2013 TCCCTAACCATTTGGGACTGGAAATGGAGGACACAACTTTGTATGCCATTTATTTATCCATG 2072  
1742 CAAACAAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATAGTGAGAGTCTTTTCCAACT 1801  
2073 CAAACAAAGTCATTAAAGAAACATTACCTGGAGCCAGAAATAGTGAGAGTCTTTTCCAACT 2132  
1802 TGTCTCTTCTTCCGAGGAATGGAGCCAAATTCGAGAGCAATGCAATGGGTTTTTCTCTT 1861  
2133 TGTCTCTTCTTCCGAGGAATGGAGCCAAATTCGAGAGCAATGCAATGGGTTTTTCTCTT 2192  
1862 ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGATGATGATGATGATGATGATGAT 1921  
2193 ACCATGCAATCAAGTCTGGCATGGACATGGAGATGATGATGATGATGATGATGATGATGAT 2252  
1922 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC 1981  
2253 ATGATGATATCCATAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAGAGACC 2312  
1982 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGGAGGAGGAGGAGGAG 2041  
2313 CTGAGGCCACTGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGGAGGAGGAGGAGGAG 2372  
2042 TCATTCAGACTGATGAGGAGAAATGGCCCTCTGAGAGACCCCTTGTAGTATGCCCTTGT 2101  
2373 TCATTCAGACTGATGAGGAGAAATGGCCCTCTGAGAGACCCCTTGTAGTATGCCCTTGT 2432  
2102 TGAAGGGCATTTGAAACATATTATTGAGGATCTGAGGAGCCAGGTTAAACCAAAAAA 2161  
2433 TGAAGGGCATTTGAAACATATTATTGAGGATCTGAGGAGCCAGGTTAAACCAAAAAA 2492  
2162 AATATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGATGATGATGATGATGATGATG 2221  
2493 AATATCCCGACCTCTCAATATATTTGAAGGACCCCTGATGATGATGATGATGATGATGATG 2552  
2222 GTGATCTTTTGGAGCTGGAAATGTTTCTACCTCAGGTATTAAGTCAAGCCCGGCTTA 2281  
2553 GTGATCTTTTGGAGCTGGAAATGTTTCTACCTCAGGTATTAAGTCAAGCCCGGCTTA 2612  
2282 TGAAGAGGCTGTTGGCCACTTATCCCTTTCATGGAAGAAAGAGAGAGAGAGAGAGAG 2341  
2613 TGAAGAGGCTGTTGGCCACTTATCCCTTTCATGGAAGAAAGAGAGAGAGAGAGAGAGAG 2672

2342 TGCCTTAAACGCGCAGTAGAAG 2401  
2673 TGCCTTAAACGCGCAGTAGAAG 2732  
2402 TTAAGAGGCGCAGCTGCACGACATAGGCAAGAAACATAGTTCGAGTAGTCTTGGCTGCAATA 2461  
2733 TTAAGAGGCGCAGCTGCACGACATAGGCAAGAAACATAGTTCGAGTAGTCTTGGCTGCAATA 2792  
2462 ATTTCGAGTATTGATTTAGGAGTCAATCTCCATGTGATTAAGATACATGAAAGCTGCTC 2521  
2793 ATTTCGAGTATTGATTTAGGAGTCAATCTCCATGTGATTAAGATACATGAAAGCTGCTC 2852  
2522 TTGACCAACAGCAGATATTAATTTGGCCTCTCAGGACTCATCTCTTCCCTGGATGAAA 2581  
2853 TTGACCAACAGCAGATATTAATTTGGCCTCTCAGGACTCATCTCTTCCCTGGATGAAA 2912  
2582 TGAATTTTGTGTCGAAGGAATTTAGGAGATTTAGCTTATAGGATTTCCATTTGTTGATTTGAG 2641  
2913 TGAATTTTGTGTCGAAGGAATTTAGGAGATTTAGCTTATAGGATTTCCATTTGTTGATTTGAG 2972  
2642 GAGCAACCACTTCAAAAAACCCACACAGCAGTGTAAATAGCTCCGAGATACAGTGACCTG 2701  
2973 GAGCAACCACTTCAAAAAACCCACACAGCAGTGTAAATAGCTCCGAGATACAGTGACCTG 3032  
2702 TGAATCCATGTCTGAGAGCTCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761  
3033 TGAATCCATGTCTGAGAGCTCAAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 3092  
2762 ATCTAAGGATCAATATTTTGGAGAAATCATGGAAGAAATATGAAGATATTTAGACAGACC 2821  
3093 ATCTAAGGATCAATATTTTGGAGAAATCATGGAAGAAATATGAAGATATTTAGACAGACC 3152  
2822 ATTATGAGTCTCTCAAGGAGAGAGATCTTACCCTTAACTCAAGCCAGGAAAGTGGTT 2881  
3153 ATTATGAGTCTCTCAAGGAGAGAGATCTTACCCTTAACTCAAGCCAGGAAAGTGGTT 3212  
2882 TCAAAATGATTTGGCTGTCTGAACTCCACAGTGAAGCCACGTTTATTTGGAGCCAGG 2941  
3213 TCAAAATGATTTGGCTGTCTGAACTCCACAGTGAAGCCACGTTTATTTGGAGCCAGG 3272  
2942 TCTTTGAGACTATGACCTGCGAGAGCTGCTGAGACTGATGAGTGAAGCCCTTTCTTTG 3001  
3273 TCTTTGAGACTATGACCTGCGAGAGCTGCTGAGACTGATGAGTGAAGCCCTTTCTTTG 3332  
3002 ATGCTCTGGCAGCTCCGGGCAAGTACCCGAATCGAGGCTTCCCCAGAGATTTTAAACGACA 3061  
3333 ATGCTCTGGCAGCTCCGGGCAAGTACCCGAATCGAGGCTTCCCAAGATTTTAAACGACA 3392  
3062 AAACAGTAGGTGGAGAGGCGCAGAGAGGTTCTAGATGATGCCCAATATGCTGGAACACAC 3121  
3393 AAACAGTAGGTGGAGAGGCGCAGAGAGGTTCTAGATGATGCCCAATATGCTGGAACACAC 3452  
3122 TGAATAGTCAAAAGAACTCCGGGCGCGGGTGTGGTGGTCTGAGAGCCCTTTCTTTG 3181  
3453 TGAATAGTCAAAAGAACTCCGGGCGCGGGTGTGGTGGTCTGAGAGCCCTTTCTTTG 3512  
3182 TCAAGAGCAGCATTTCACTGTATCGCAGAGAGCTGCTGTGCCCGGCTGCGAGAGCCCATAG 3241  
3513 TCAAGAGCAGCATTTCACTGTATCGCAGAGAGCTGCTGTGCCCGGCTGCGAGAGCCCATAG 3572  
3242 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGAGCTCTGCCAGACCGAGCCATCT 3301  
3573 CCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGAGCTCTGCCAGACCGAGCCATCT 3632  
3302 ACTGCTCTCAGACTTCATGCTCTCCCTTGCATTTCTGGCATTCCTGACTACCTTGGGCTGT 3361  
3633 ACTGCTCTCAGACTTCATGCTCTCCCTTGCATTTCTGGCATTCCTGACTACCTTGGGCTGT 3692  
3362 TTGCCGCTTCTGCTGCTTTGGGGTAGAAGAGCTCAGCAAGCCCTTATGAGGATGATGGAGC 3421  
3693 TTGCCGCTTCTGCTGCTTTGGGGTAGAAGAGCTCAGCAAGCCCTTATGAGGATGATGGAGC 3752  
3422 ACTACAGCAGCATCATGTTGTTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCGCTTTGCAAGAG 3481

```
Db 3753 ACTACAGCAGCATCATGGTCAAGCGCTGGGAGACCGGCTGGCAGAGGCGCTTTCAGAGAG 3812
Qy 3482 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGCCCTACTGTGGCAGTGTGAGCAGCTGGAGC 3541
Db 3813 AGCTCCATGAAGAGTTGCGCGAGAACTGTGGCCCTACTGTGGCAGTGTGAGCAGCTGGAGC 3872
Qy 3542 TCCAGACCTGCGAAGTTGCGGTACAGGGCATCCGCCGCTCTGGCTACCCAGCC 3601
Db 3873 TCCAGACCTGCGCAGGCTGCGGTACAGGGCATCCGCCGCTCTGGCTACCCAGCC 3932
Qy 3602 AGCCCGACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGACAGCTCTACAG 3661
Db 3933 AGCCCGACACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGACAGCTCTACAG 3992
Qy 3662 GCATTAGGTTAAACAGAACTATTAGCAATGGCACCTGCTCAGCAGTCTCAGGCTCTACT 3721
Db 3993 GCATTAGGTTAAACAGAACTATTAGCAATGGCACCTGCTCAGCAGTCTCAGGCTCTACT 4052
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTCTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGG 3781
Db 4053 TCTCCAAATTTGAAGTCCAAATATTTTCTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGG 4112
Qy 3782 ATTATGCAATTTGAGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCCATTT 3841
Db 4113 ATTATGCAATTTGAGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGGACCCCATTT 4172
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTT 3875
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTT 4206
```

## RESULT 7

```
US-09-962-665-1
; Sequence 1, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 194, 3209
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1136, 1334, 3150, 5551, 5934
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 284, 1252, 1699, 5573, 5659, 5678, 5874
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 3207
; OTHER INFORMATION: n = g or t
```

```
; NAME/KEY: misc feature
; LOCATION: 5444
; OTHER INFORMATION: n = c or a
US-09-962-665-1
```

```
Query Match 79.3%; Score 3109; DB 3; Length 7224;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3859; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
Qy 2 GTACCTGTGGAGAGACGCTCTTCTCTGCGGCGCCTCTGCGCAAGGAGAGACTGAC 61
Db 333 GTACCTGTGGAGAGACGCTCTTCTCTGCGGCGCCTCTGCGCAAGGAGAGACTGAC 392
Qy 62 ACATGTCCACCCCGCTCCAAAGACCTGTCCAAACCGAAGGTCTGAAGAAAAACCTCGGG 121
Db 393 ACATGTCCACCCCGCTCCAAAGACCTGTCCAAACCGAAGGTCTGAAGAAAAACCTCGGG 452
Qy 122 ATGAGATCAATGCCATTTCTGCAGAGAGAGGATTTATGCTGCTGGATGGAGGATGGGACCA 181
Db 453 ATGAGATCAATGCCATTTCTGCAGAGAGAGGATTTATGCTGCTGGATGGAGGATGGGACCA 512
Qy 182 TGATCCAGCGGGAGAGAGCTTAAACGAAAGAACACTTCCGAGGTGAGAAATTTAAAGATCATG 241
Db 513 TGATCCAGCGGGAGAGAGCTTAAACGAAAGAACACTTCCGAGGTGAGAAATTTAAAGATCATG 572
Qy 242 CCAGGCGCTGAAAGGCAACATTTTAAAGTATACTCAGCCCTGATGTCATTTACC 301
Db 573 CCAGGCGCTGAAAGGCAACATTTTAAAGTATACTCAGCCCTGATGTCATTTACC 632
Qy 302 AAATCCATAAGGAATACTTGTGCTGGCTGGGCGCAGATATCATTTGAAACAAATACTTTTAGCA 361
Db 633 AAATCCATAAGGAATACTTGTGCTGGCTGGGCGCAGATATCATTTGAAACAAATACTTTTAGCA 692
Qy 362 GCATAGTATGCCCAGCTGACTATGGCTTGAACACTTGGCTACCGATGAACATGT 421
Db 693 GCATAGTATGCCCAGCTGACTATGGCTTGAACACTTGGCTACCGATGAACATGT 752
Qy 422 GCTCTCGAGGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 481
Db 753 GCTCTCGAGGAGTGGCCAGAAAGCTGCGAGGAGGTAACTCTCCAGACAGGAATTAAGA 812
Qy 482 GGTCTTGTGCGAGGCGCTCTGGGTCCGACTAAATAGACACACTCTGTGTCCCACTGTGG 541
Db 813 GGTCTTGTGCGAGGCGCTCTGGGTCCGACTAAATAGACACACTCTGTGTCCCACTGTGG 872
Qy 542 AAAGCGCGGATATAGGAACATCAATTTGATGAGCTTGTGAGCATATCAAGAGCAGG 601
Db 873 AAAGCGCGGATATAGGAACATCAATTTGATGAGCTTGTGAGCATATCAAGAGCAGG 932
Qy 602 CCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTGAAACTATTTTGTGATACTG 661
Db 933 CCAAGGACTTCTGGATGGCGGGTTGATATCTTACTCATTGAAACTATTTTGTGATACTG 992
Qy 662 CCAATGCCAAGCAGCGCTTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 721
Db 993 CCAATGCCAAGCAGCGCTTGTGCACTCCAAATCTTTTGGAGGAGAAATATGCTCCCC 1052
Qy 722 GGCTATCTTTATTTTCAGGAGCGATCGTTGATATAAGTGGGCGGACTCTTTCCGAGCAGA 781
Db 1053 GGCTATCTTTATTTTCAGGAGCGATCGTTGATATAAGTGGGCGGACTCTTTCCGAGCAGA 1112
Qy 782 CAGGAGGAGGATTTGTTCATCAGCGGTCTCTACGAGAACCACTCTGCTGATTTGATTAATT 841
Db 1113 CAGGAGGAGGATTTGTTCATCAGCGGTCTCTACGAGAACCACTCTGCTGATTTGATTAATT 1172
Qy 842 GTGCTTTGGGTGCGAGCTGAGATGAGACTTTTATTTGAAATAATTTGAAAAATGTACACAG 901
Db 1173 GTGCTTTGGGTGCGAGCTGAGATGAGACTTTTATTTGAAATAATTTGAAAAATGTACACAG 1232
Qy 902 CCTATGTCTCTGTGTTATCCCAATGCGAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 961
Db 1233 CCTATGTCTCTGTGTTATCCCAATGCGAGGTCTTCCCAACACCTTTGGTGACTATGATGAAA 1292
```





Db 3453 TGATTAGTCAAAAGAAACTCGGGCCCGGGTGTGGTTGGTTCTGGCCAGCACAGAGTA 3512  
Qy 3182 TCCAAGCAGCATTCACCTGTACGACAGAGCTGCTGTGCCCGAGGCTGCAGAGCCCATAG 3241  
Db 3513 TCCAAGCAGCATTCACCTGTACGCGGAGCTGCTGTGCCCGAGGCTGCAGAGCCCATAG 3572  
Qy 3242 CCACCTTCTATGGGTAAAGCAACAGGCTCAGAGGACTCTGCCAGCACGGAGCCATACT 3301  
Db 3573 CCACCTTCTATGGGTAAAGCAACAGGCTCAGAGGACTCTGCCAGCACGGAGCCATACT 3632  
Qy 3302 ACTGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGCGATCCGTAATCTGCGGCTGT 3361  
Db 3633 ACTGCCTCTCAGACTTCATCGCTCCCTTGCAATCTGCGATCCGTAATCTGCGGCTGT 3692  
Qy 3362 TTGCCGTGCTGCTTTGGGGTAGAAGAGCTGAGCAGGCTATGAGGATGATGGTGAAG 3421  
Db 3693 TTGCCGTGCTGCTTTGGGGTAGAAGAGCTGAGCAGGCTATGAGGATGATGGTGAAG 3752  
Qy 3422 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAGAG 3481  
Db 3753 ACTACAGCAGCATCATGGTCAAGGCGCTGGGGGACCGGCTGGCAGAGGCCCTTTGCAGAG 3812  
Qy 3482 AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGAAG 3541  
Db 3813 AGCTCCATGAAGAGTTCCGCGAGAACTGTGGGCTACTGTGGCAGTGAGCAGCTGAAG 3872  
Qy 3542 TCGCAGACCTGCGAGAGTTGGGTTACAGGGCATCCGCGGCTCTGGCTACCCAGCC 3601  
Db 3873 TCGCAGACCTGCGAGAGTTGGGTTACAGGGCATCCGCGGCTCTGGCTACCCAGCC 3932  
Qy 3602 AGCCCGACCAACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661  
Db 3933 AGCCCGACCAACACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3992  
Qy 3662 GCATTAGGTTAAACAGAAATCAATTAGCAATGGCAGCTGCTTCAGCAGTCTCAGGCTCTACT 3721  
Db 3993 GCATTAGGTTAAACAGAAATCAATTAGCAATGGCAGCTGCTTCAGCAGTCTCAGGCTCTACT 4052  
Qy 3722 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 3781  
Db 4053 TCTCCAAATTTGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 4112  
Qy 3782 ATATGCAATTCAGGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGAGCCCATTT 3841  
Db 4113 ATATGCAATTCAGGAGAAACATATCTGTGGCTGAGGTTGAGAAATGGCTTGAGCCCATTT 4172  
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTT 3875  
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTT 4206

## RESULT 8

US-09-963-333-1  
; Sequence 1, Application US/09963333  
; Patent No. 6664062  
; GENERAL INFORMATION:  
; APPLICANT: Staton, Jr., Vincent P.  
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES  
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT  
; TITLE OF INVENTION: OF DISEASE  
; FILE REFERENCE: 11926-015002  
; CURRENT APPLICATION NUMBER: US/09/963,333  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/658,659  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 09/596,033  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 09/357,743  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 09/357,024  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: 60/093,484  
; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 7224  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 194..3209  
; OTHER INFORMATION: n = c or g  
; NAME/KEY: misc feature  
; LOCATION: 1136, 1334, 3150, 5551, 5934  
; OTHER INFORMATION: n = a or g  
; NAME/KEY: misc feature  
; LOCATION: 284..1252, 1699, 5573, 5659, 5678, 5874  
; OTHER INFORMATION: n = c or t  
; NAME/KEY: misc feature  
; LOCATION: 3207  
; OTHER INFORMATION: n = g or t  
; NAME/KEY: misc feature  
; LOCATION: 5444  
; OTHER INFORMATION: n = c or a  
; US-09-963-333-1  
  
Query Match 79.3%; Score 3109; DB 3; Length 7224;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 3859; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
  
Qy 2 GTCACTGTGAGAGACAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 61  
Db 333 GTCACTGTGAGAGACAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGAGACTCGACA 392  
Qy 62 ACATGTCAACCGCGCTCCAGAGCTCTGCGCAACCGCAAGCTCTGAAGAAACCCCTCGGG 121  
Db 393 ACATGTCAACCGCGCTCCAGAGCTCTGCGCAACCGCAAGCTCTGAAGAAACCCCTCGGG 452  
Qy 122 ATCAGATCAATGTCATTTCTGCAAGAGGATTAATGCTGCTGGAATGGAGGATGGGACCA 181  
Db 453 ATGAGATCAATGTCATTTCTGCAAGAGGATTAATGCTGCTGGAATGGAGGATGGGACCA 512  
Qy 182 TGATCCAGCGGGAAGAGCTTAAACGAGAACACTTCCGAGGTCAAGAAATTTAAAGATCATG 241  
Db 513 TGATCCAGCGGGAAGAGCTTAAACGAGAACACTTCCGAGGTCAAGAAATTTAAAGATCATG 572  
Qy 242 CCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTAACTCAGCCTGATGTCATTTACC 301  
Db 573 CCAGGCGCTGAAAGGCAACAATGACATTTTAAAGTAACTCAGCCTGATGTCATTTACC 632  
Qy 302 AAATCCATAAGGAATACTTCTGCTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA 361  
Db 633 AAATCCATAAGGAATACTTCTGCTGGCTGGGCGAGATATCATTTGAAACAAATACCTTTAGCA 692  
Qy 362 GCATAGTATTTGCCAAGCTGACTATGGCCCTTGAACACTTTGGCCCTACCGATGACATGT 421  
Db 693 GCATAGTATTTGCCAAGCTGACTATGGCCCTTGAACACTTTGGCCCTACCGATGACATGT 752  
Qy 422 GCTCTCAGAGGTGGCCAGAAAAGCTCCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTCAGAGGTGGCCAGAAAAGCTCCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGTGGCAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG 541  
Db 813 GGTTTGTGGCAGGGGCTCTGGGTCGACTAATAAGACACTCTCTGTGTCCCATCTGTGG 872  
Qy 542 AAAGCCCGGATTAAGGAAACATCACTTTGAGAGCTTTGTAAGACATACCAAGAGCAGG 601  
Db 873 AAAGCCCGGATTAAGGAAACATCACTTTGAGAGCTTTGTAAGACATACCAAGAGCAGG 932  
Qy 602 CCAAGGACTTCTGGATGGGGGTTGATCTTACTCATTGAACTATTTTTCATCTG 661  
Db 933 CCAAGGACTTCTGGATGGGGGTTGATCTTACTCATTGAACTATTTTTCATCTG 992  
Qy 662 CCAATGCCAAGGAGCCTTGTGTGCACTCCAAATCTTTTGTGAGGAGAAATATGCTCCC 721



Db 993 CCAATGCAAGGAGCCTTGTGTGCACTCCAAATCTTTTGGAGGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTTCAAGGACAGATCGTTGATAAAGTGGCGGAGCTCTTTCCGGACAGA 781  
Db 1053 GGCCTATCTTTATTTCAAGGACAGATCGTTGATAAAGTGGCGGAGCTCTTTCCGGACAGA 1112  
Qy 782 CAGGAGGAGATTTGTCATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGATTTAAAT 841  
Db 1113 CAGGAGGAGATTTGTCATCAGCGTGTCTCATGAGAACCACTCTGCAATTTGATTTAAAT 1172  
Qy 842 GTGCTTTGGGTGCACTGAGATGAGACCTTTTATGAAATAATTTGGAATAATGACACAG 901  
Db 1173 GTGCTTTGGGTGCACTGAGATGAGACCTTTTATGAAATAATTTGGAATAATGACACAG 1232  
Qy 902 CCTATGTCCTCTGTTATCCCAATGCAAGGTCTTCCCAACACTTTTGGTGACTATGATGAAA 961  
Db 1233 CCTATGTCCTCTGTTATCCCAATGCAAGGTCTTCCCAACACTTTTGGTGACTATGATGAAA 1292  
Qy 962 CGCCTTCTATGATGGCCCAAGCACTAAAGATTTTGTATGATGGCTTGGTCAATATAG 1021  
Db 1293 CGCCTTCTATGATGGCCCAAGCACTAAAGATTTTGTATGATGGCTTGGTCAATATAG 1352  
Qy 1022 TTGAGGATGCTGTGGGTCAACCAACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGAGGATGCTGTGGGTCAACCAACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTGTAAGCCTTAGAGTTCCACCTGCCACTCTCTTTGAAGGACATATGTTACTCTGTGGTC 1141  
Db 1413 ATTGTAAGCCTTAGAGTTCCACCTGCCACTCTCTTTGAAGGACATATGTTACTCTGTGGTC 1472  
Qy 1142 TAGAGCCCTTCAGATTTGGAACCGTACACCAACTTTTGTATACTGAGAGCGCTGTAATG 1201  
Db 1473 TAGAGCCCTTCAGATTTGGAACCGTACACCAACTTTTGTATACTGAGAGCGCTGTAATG 1532  
Qy 1202 TTGAGGATCAAGGATTTGCTTAACTCATATGCGAGAACTATGAGAGCCTTGT 1261  
Db 1533 TTGAGGATCAAGGATTTGCTTAACTCATATGCGAGAACTATGAGAGCCTTGT 1592  
Qy 1262 GTGTTGCAAGTTCAGAGTGGAAATGGAGCCAGGTGTTGGATGTCAACATGGATGATG 1321  
Db 1593 GTGTTGCAAGTTCAGAGTGGAAATGGAGCCAGGTGTTGGATGTCAACATGGATGATG 1652  
Qy 1322 GCATGCTAGATGTCGAAGTGCATATGACCAAGATTTTGCACACTTAAATGCTTCCGAGCCAG 1381  
Db 1653 GCATGCTAGATGTCGAAGTGCATATGACCAAGATTTTGCACACTTAAATGCTTCCGAGCCAG 1712  
Qy 1382 ACATCGCAAGGTACCTTTGTGTCATCGACTCCTCCCAATTTTGTCTGATGATGAGCTGGGT 1441  
Db 1713 ACATCGCAAGGTACCTTTGTGTCATCGACTCCTCCCAATTTTGTCTGATGATGAGCTGGGT 1772  
Qy 1442 TAAAGTGTCTGCCAAGGAAAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAAGAGAGACG 1501  
Db 1773 TAAAGTGTCTGCCAAGGAAAGTGCATTTGTCAATAGCAATTTAGTCTGAAGGAAGAGAGACG 1832  
Qy 1502 ACTTCTTGGAGAGGCGCAGGAAGATTAAGATGATGAGCTGCTATGTTGGTTCATGCTT 1561  
Db 1833 ACTTCTTGGAGAGGCGCAGGAAGATTAAGATGATGAGCTGCTATGTTGGTTCATGCTT 1892  
Qy 1562 TTGATGAAGAGAGACAGGCAAGAGAAACAGACACAAATTCAGAGTGTGACCCGGGCT 1621  
Db 1893 TTGATGAAGAGAGACAGGCAAGAGAAACAGACACAAATTCAGAGTGTGACCCGGGCT 1952  
Qy 1622 ACCATCTGCTTGTGAAATACTGGGCTTTAATCAATGACATTTATTTTGGCCCTTAATA 1681  
Db 1953 ACCATCTGCTTGTGAAATACTGGGCTTTAATCAATGACATTTATTTTGGCCCTTAATA 2012  
Qy 1682 TCCTAACCAATTTGGGACTGGAAATGGAGAAACAACTTTGATGCCATTAATTTTATCCATG 1741  
Db 2013 TCCTAACCAATTTGGGACTGGAAATGGAGAAACAACTTTGATGCCATTAATTTTATCCATG 2072  
Qy 1742 CAAACAAAGTCATTAAGAAACATTAACCTGGAGCGAGATTAAGTGGAGGTCTTTTCCAACT 1801

Db 2073 CAAACAAAGTCATTAAGAAACATTAACCTGGAGCGAGATTAAGTGGAGGTCTTTTCCAACT 2132  
Qy 1802 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAAGCAATGCAATGCGGTTTTTCTCTTT 1861  
Db 2133 TGTCTCTTCTCTTCCGAGGAATGGAAGCCATTCGAGAAGCAATGCAATGCGGTTTTTCTCTTT 2192  
Qy 1862 ACCATGCAATCAAGTCTGCGATGGAATGGAAGTGTGATGCTGGAACCTTCCCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGCGATGGAATGGAAGTGTGATGCTGGAACCTTCCCTGTGT 2252  
Qy 1922 ATGATCATATCCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db 2253 ATGATCATATCCATTAAGGAACCTTCTGAGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
Qy 1982 CTGAGGCCACTGAGAAAGCTCTTAAGTTATGCCAGACTCAAGGCAAGAGGGAAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAAAGCTCTTAAGTTATGCCAGACTCAAGGCAAGAGGGAAGAAAG 2372  
Qy 2042 TCATTCAGACTGATGATGGAAGATTTGGCCCTGTGCAAGAACGCTTGTGATGATGCCCTTG 2101  
Db 2373 TCATTCAGACTGATGATGGAAGATTTGGCCCTGTGCAAGAACGCTTGTGATGATGCCCTTG 2432  
Qy 2102 TGAAGGGCAATTCGAAAAACATATTTATGAGATATCTGAGGAAGCCAGGTTAAACCAAAAA 2161  
Db 2433 TGAAGGGCAATTCGAAAAACATATTTATGAGATATCTGAGGAAGCCAGGTTAAACCAAAAA 2492  
Qy 2162 AATATCCCCGACCTCTCAATATATAATTTGAAGGACCCCTGATGATGGAATGGAATTTGTTG 2221  
Db 2493 AATATCCCCGACCTCTCAATATATAATTTGAAGGACCCCTGATGATGGAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAAGATTTTCTACTCTCAGGTTATAAGTCAGCCCGGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAAGATTTTCTACTCTCAGGTTATAAGTCAGCCCGGGTTA 2612  
Qy 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAAACCCAGAG 2341  
Db 2613 TGAAGAGGCTGTGGCCACCTTATCCCTTTCATGGAAGAAAGAGAGAAACCCAGAG 2672  
Qy 2342 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACAGGGCCACCATCTGCTGGCCACTG 2401  
Db 2673 TGCTTAAACGGCACAGTAGAAGAGAGGACCCCTTACAGGGCCACCATCTGCTGGCCACTG 2732  
Qy 2402 TTAAGGGCGACGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2461  
Db 2733 TTAAGGGCGACGTGCACGACATAGGCAAGAACATAGTTGGAGTAGTCTTTGGCTGCAATA 2792  
Qy 2462 ATTTCCGAGTATTTGATTTAGGAGTCATGCTCCATGTCATGATGATGATGATGATGCTC 2521  
Db 2793 ATTTCCGAGTATTTGATTTAGGAGTCATGCTCCATGTCATGATGATGATGATGATGCTC 2852  
Qy 2522 TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTCTCCCTGGATGAAA 2581  
Db 2853 TTGACCAAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTCTCCCTGGATGAAA 2912  
Qy 2582 TGATTTTGTGTCGAAGGAATGGAAGATTTAGCTATAAGGATTTCCATTTGTTGATTTGGAG 2641  
Db 2913 TGATTTTGTGTCGAAGGAATGGAAGATTTAGCTATAAGGATTTCCATTTGTTGATTTGGAG 2972  
Qy 2642 GAGCAACCACTTCAAAACCCACAGCAGTGTAAATAGCTCCGAGATACAGTGCACCTG 2701  
Db 2973 GAGCAACCACTTCAAAACCCACAGCAGTGTAAATAGCTCCGAGATACAGTGCACCTG 3032  
Qy 2702 TAAATCATGCTCTGAGAGCGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 2761  
Db 3033 TAAATCATGCTCTGAGAGCGCTCCAGAGTGTGGTGTGTTCCAGCTGTTAGATGAAA 3092  
Qy 2762 ATCTAAGGATGAATATCTTTGAGGAAATCATCGAAGAAATATGAAGATATTAGCAGGACC 2821  
Db 3093 ATCTAAGGATGAATATCTTTGAGGAAATCATCGAAGAAATATGAAGATATTAGCAGGACC 3152  
Qy 2822 ATTATGAGTCTCTCAAGGAGAGAGATATCTTACCTTTAAGTCAAGCCAGAAAGAGTGT 2881  
Db 3153 ATTATGAGTCTCTCAAGGAGAGAGATATCTTACCTTTAAGTCAAGCCAGAAAGAGTGT 3212

```
Qy 2882 TCCAAATGGATTGGCTGCTCTGAACCTCACCAGTGAAGGCCACGTTTATTGGGACCAGG 2941
Db 3213 TCCAAATGGATTGGCTGCTCTGAACCTCACCAGTGAAGGCCACGTTTATTGGGACCAGG 3272
Qy 2942 TCTTTGAAGACTATGACTTCGAGAGCTGGTGGACTACATTGACTGGAAGCCTTTCTTTG 3001
Db 3273 TCTTTGAAGACTATGACTTCGAGAGCTGGTGGACTACATTGACTGGAAGCCTTTCTTTG 3332
Qy 3002 ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGCTTCCCGAAGATATTTAAAGACA 3061
Db 3333 ATGTCTGGCAGCTCCGGGGCAAGTACCCGAATCGAGCTTCCCGAAGATATTTAAAGACA 3392
Qy 3062 AAACAGTAGTGGAGAGCCAGAGAGCTCTACGATGATGCCCAATATGCTGAACACAC 3121
Db 3393 AAACAGTAGTGGAGAGCCAGAGAGGTCTACGATGATGCCCAATATGCTGAACACAC 3452
Qy 3122 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3181
Db 3453 TGATTAGTCAAAAGAACTCCGGGCCCGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3512
Qy 3182 TCCAAAGCAGATTCACTGTATCGCAGAGGTGCTGTGCCCCCAGGCTGCGAGGCCCATAG 3241
Db 3513 TCCAAAGCAGATTCACTGTATCGCAGAGGTGCTGTGCCCCCAGGCTGCGAGGCCCATAG 3572
Qy 3242 CCCTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACGGGCCATCT 3301
Db 3573 CCCTTTCTATGGTTAAGGCAACAGGCTGAGAGGACTCTGCCAGCACGGGCCATCT 3632
Qy 3302 ACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCTGGCATCCGCTGACTTCTGGCCTGT 3361
Db 3633 ACTGCTCTCAGACTTTCATCGCTCCCTTGCATTTCTGGCATCCGCTGACTTCTGGCCTGT 3692
Qy 3362 TTGCGGTTGCTGCTTTGGGGTAGAAGCTGAGCAAGGCTATGAGGATGATGGTACG 3421
Db 3693 TTGCGGTTGCTGCTTTGGGGTAGAAGCTGAGCAAGGCTATGAGGATGATGGTACG 3752
Qy 3422 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAAG 3481
Db 3753 ACTACAGCAGCATCATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAAG 3812
Qy 3482 AGCTCCATGAAGAGTTTCGCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG 3541
Db 3813 AGCTCCATGAAGAGTTTCGCGAGAACTGTGGGCTTACTGTGGCAGTGAGCAGCTGGACG 3872
Qy 3542 TCGCAGACTCGGAAGGTTGGGTTACAAAGGCTATCCGCGCGGCTCCTGGCTACCCAGCC 3601
Db 3873 TCGCAGACTCGGCGAGGCTGGGTTACAAAGGCTATCCGCGCGGCTCCTGGCTACCCAGCC 3932
Qy 3602 AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTCGCAGACATCGAGCAGTCTACAG 3661
Db 3933 AGCCCGACCAACCGAGAGCTCACCATGTGGAGACTTTGCAGACATCGAGCAGTCTACAG 3992
Qy 3662 GCATTAGGTTAAAGAAATCAATTAGCAATGGCACTGTGCTCAGCAGTCTCAGGCTCTACT 3721
Db 3993 GCATTAGGTTAAAGAAATCAATTAGCAATGGCACTGTGCTCAGCAGTCTCAGGCTCTACT 4052
Qy 3722 TCTCCAAATTCGAAGTCCAAATATTTTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 3781
Db 4053 TCTCCAAATTCGAAGTCCAAATATTTTGTGTGTGGGGAAGATTTCCAAAGATCAGGTTGAGG 4112
Qy 3782 ATTTATGATTCAGGAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGAGCCCATTT 3841
Db 4113 ATTTATGATTCAGGAAGCAATATCTGTGGCTGAGGTTGAGAAATGGCTTGAGCCCATTT 4172
Qy 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3875
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 4206
```

RESULT 9

US-09-962-677-1

; Sequence 1, Application US/09962677

```
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: 194..3209
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc feature
; LOCATION: 1136..1334, 3150..5551, 5934
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc feature
; LOCATION: 284..1252, 1699..5573, 5659..5678, 5874
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc feature
; LOCATION: 3207
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc feature
; LOCATION: 5444
; OTHER INFORMATION: n = c or a
; US-09-962-677-1
```

```
Query Match 79.3%; Score 3109; DB 3; Length 7224;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3859; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 2 GTCACTGTGTGAGAGCAGCTCTTCTCTGCGCGGCCCTCTGCGCAAGGAGAGACTCGACA 61
Db 333 GTCACTGTGTGAGAGCAGCTCTTCTCTGCGCGGCCCTCTGCGCAAGGAGAGACTCGACA 392
Qy 62 ACATGTCAACCGCGCTCCAAAGACTGTGCGCAACCCGAGGTCTGAAGAAACCTTCGGG 121
Db 393 ACATGTCAACCGCGCTCCAAAGACTGTGCGCAACCCGAGGTCTGAAGAAACCTTCGGG 452
Qy 122 ATGAGATCAATGCTTCTGCAAGAGGATTAATGCTGCGGATGAGGATGCGGACCA 181
Db 453 ATGAGATCAATGCTTCTGCAAGAGGATTAATGCTGCGGATGAGGATGCGGACCA 512
Qy 182 TGATCCAGCGGGAGAGACTAAACGAAGAACACTTCCGAGGTCAAGGATTTAAAGATCATG 241
Db 513 TGATCCAGCGGGAGAGACTAAACGAAGAACACTTCCGAGGTCAAGGATTTAAAGATCATG 572
Qy 242 CCAGCGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGCTATTACC 301
Db 573 CCAGCGCGCTGAAAGGCAACATGACATTTTAAGTATACTCAGCCTGATGCTATTACC 632
Qy 302 AAATCCATAAGGAATACCTGCTGCGGGCAGATATCAATTGAAACAAATACCTTTAGCA 361
Db 633 AAATCCATAAGGAATACCTGCTGCGGGCAGATATCAATTGAAACAAATACCTTTAGCA 692
Qy 362 GCCTAGTATTTGCCCAAGCTGACTATGGCTTGAACACTTGGGCTACCGGATGAACATGT 421
```

Db 693 GCACTAGTATTGCCAAGCTGACTATGGCCCTTGAACTTTGGCCCTACCGGATGAACTATGT 752  
Qy 422 GCTCTGAGAGTGGCCAGAAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 481  
Db 753 GCTCTGAGAGTGGCCAGAAAAGCTGCCGAGAGGTAACTCTCCAGACAGGAATTAAGA 812  
Qy 482 GGTTTGGGAGGGGCTCTGGGTCCGACTAATGAACACTCTCTGTGTCCCACTCTGTGG 541  
Db 813 GGTTTGGGAGGGGCTCTGGGTCCGACTAATGAACACTCTCTGTGTCCCACTCTGTGG 872  
Qy 542 AAAGGCCGGATATAGGAACATCACTTTGATGAGCTTTGTAAGCATATCAAGAGCAGG 601  
Db 873 AAAGGCCGGATATAGGAACATCACTTTGATGAGCTTTGTAAGCATATCAAGAGCAGG 932  
Qy 602 CCAAAGGACTCTTGATGGCGGGTTGATATCTTACTCATATGAAACTATTTTGTGATCTG 661  
Db 933 CCAAAGGACTCTTGATGGCGGGTTGATATCTTACTCATATGAAACTATTTTGTGATCTG 992  
Qy 662 CCAATGCCAAGGAGCGCTTTGCACTCCAAATCTTTTGGAGGAAATATGCTCCCC 721  
Db 993 CCAATGCCAAGGAGCGCTTTGCACTCCAAATCTTTTGGAGGAAATATGCTCCCC 1052  
Qy 722 GGCCTATCTTTATTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGACAGA 781  
Db 1053 GGCCTATCTTTATTCAGGACGATCGTTGATATAAAGTGGCGGACTCTTTCCGACAGA 1112  
Qy 782 CAGGAGGAGATTGTGCATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAT 841  
Db 1113 CAGGAGGAGATTGTGCATCAGCGTGTCTCATGGAGAACCACTCTGCATTTGGATTAAT 1172  
Qy 842 GTGCTTTGGGTGACGTGAGATGAGACCTTTTATTTGAAATTAATTTGAAATTTACAACAG 901  
Db 1173 GTGCTTTGGGTGACGTGAAATGAGACCTTTTATTTGAAATTAATTTGAAATTTACAACAG 1232  
Qy 902 CCTATGCTCTGTATTATCCCAATGTCAGGTCTTCCCAACACTTTGGTGACTATGATGAAA 961  
Db 1233 CCTATGCTCTGTATTATCCCAATGTCAGGTCTTCCCAACACTTTGGTGACTATGATGAAA 1292  
Qy 962 CGCCTTCTATGATGGCCAGCACCTAAAGGATTTGCTATGGATGGCTGGTCAATATAG 1021  
Db 1293 CGCCTTCTATGATGGCCAGCACCTAAAGGATTTGCTATGGATGGCTGGTCAATATAG 1352  
Qy 1022 TTGGAGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1081  
Db 1353 TTGGAGATGCTGTGGTCAACACAGATCATATCAGGGAATTTGCTGAAGCTGTGAAA 1412  
Qy 1082 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTC 1141  
Db 1413 ATTGTAAAGCTTAGAGTTCCACCTGCCACTGCTTTTGAAGGACATATGTTACTGTCTGCTC 1472  
Qy 1142 TAGAGCCCTTTAGGATTTGGACCGTACACCAACTTTTGAATTTGATTTGAGAGCGCTGTAATG 1201  
Db 1473 TAGAGCCCTTTAGGATTTGGACCGTACACCAACTTTTGAATTTGATTTGAGAGCGCTGTAATG 1532  
Qy 1202 TTGCAGATCAAGAAAGTTTGTAACTCATATCATGCGAGAACTATGAAAGAGCCCTTGT 1261  
Db 1533 TTGCAGATCAAGAAAGTTTGTAACTCATATCATGCGAGAACTATGAAAGAGCCCTTGT 1592  
Qy 1262 GTGTTGCCAAAGTGCAGGTGGAATGGAGCCAGGTTTGGATGTCAACATGATGATG 1321  
Db 1593 GTGTTGCCAAAGTGCAGGTGGAATGGAGCCAGGTTTGGATGTCAACATGATGATG 1652  
Qy 1322 GCATGCTAGATGTTGCCAAGTGCATGACCAAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1381  
Db 1653 GCATGCTAGATGTTGCCAAGTGCATGACCAAGATTTTGCACCTTAATTTGCTTCCGAGCCAG 1712  
Qy 1382 ACATGCGAAAGTACCTTTGTGATGCACTCTCCCAATTTTGTGATTAAGAGCTGGGT 1441  
Db 1713 ACATGCGAAAGTACCTTTGTGATGCACTCTCCCAATTTTGTGATTAAGAGCTGGGT 1772  
Qy 1442 TAAAGTCTGCCAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAGGAGGACG 1501  
Db 1773 TAAAGTCTGCCAAGGAGTGCATTTGTCAATAGCATTAGTCTGGAAGGAGGAGGACG 1832

Qy 1502 ACTTCTTTGGAGAGGCGCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTGGTCAATGGCTT 1561  
Db 1833 ACTTCTTTGGAGAGGCGCAGGAAGATTAAAAAGTATGGAGCTGCTATGGTGGTCAATGGCTT 1892  
Qy 1562 TTGATGAGAGAGGAGCAGGCAACAGAAACAGACAAAAATCAGAGTGTGCACCCGGGCT 1621  
Db 1893 TTGATGAGAGAGGAGCAGGCAACAGAAACAGACAAAAATCAGAGTGTGCACCCGGGCT 1952  
Qy 1622 ACCATCTGCTTTGTAAGAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCCCTAATA 1681  
Db 1953 ACCATCTGCTTTGTAAGAAAACTGGGCTTTAATCCAAATGACATTAATTTTGGACCCCTAATA 2012  
Qy 1682 TCCATAACATTTGGAGCTGGAATGGAGGAACAACATTTGTATGCGCATTAATTTTATCCATG 1741  
Db 2013 TCCATAACATTTGGAGCTGGAATGGAGGAACAACATTTGTATGCGCATTAATTTTATCCATG 2072  
Qy 1742 CAAACAAAGTCATTAAAGAAACATTACTGGAGCCAGGAATAAGTGAGAGTCTTTTCCAACT 1801  
Db 2073 CAAACAAAGTCATTAAAGAAACATTACTGGAGCCAGGAATAAGTGAGAGTCTTTTCCAACT 2132  
Qy 1802 TGTCTCTTCTCTTCCGAGGAATGGAGCCATTCGAGAAGCAATGCAATGGGTTTTTCTCTT 1861  
Db 2133 TGTCTCTTCTCTTCCGAGGAATGGAGCCATTCGAGAAGCAATGCAATGGGTTTTTCTCTT 2192  
Qy 1862 ACCATGCAATCAAGTCTGCGATGGACATGGAGATAGTGAATGCTGGAACCTCCTCTGTGT 1921  
Db 2193 ACCATGCAATCAAGTCTGCGATGGACATGGGGAATAGTGAATGCTGGAACCTCCTCTGTGT 2252  
Qy 1922 ATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 1981  
Db 2253 ATGATGATATCCATAAGGAACCTTCTGACGCTCTGTGAAGATCTCATCTGGAATAAAGACC 2312  
Qy 1982 CTGAGGCCACTGAGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGGAAGAAAG 2041  
Db 2313 CTGAGGCCACTGAGAGAGCTCTTACGTTATGCCAGACTCAAGGCAAGAGGGAAGAAAG 2372  
Qy 2042 TCAATTGAGATGATGAGTGGAGAAATGGCCCTGTGCAAGAAACGCTTGTAGTATGCCCTTG 2101  
Db 2373 TCAATTGAGATGATGAGTGGAGAAATGGCCCTGTGCAAGAAACGCTTGTAGTATGCCCTTG 2432  
Qy 2102 TGAAGGGCATTCGAAAACATATTTATTTGAGGATCTGAGAGCCCTGATGAAATGAAATTTGTTG 2161  
Db 2433 TGAAGGGCATTCGAAAACATATTTATTTGAGGATCTGAGAGCCCTGATGAAATGAAATTTGTTG 2492  
Qy 2162 AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGAAATTTGTTG 2221  
Db 2493 AATATCCCGACCTCTCAATATAATTTGAAGGACCCCTGATGAAATGAAATTTGTTG 2552  
Qy 2222 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAAGTCAAGCCCGGTTA 2281  
Db 2553 GTGATCTTTTGGAGCTGGAATAATGTTTCTACCTCAGGTTATAAAGTCAAGCCCGGTTA 2612  
Qy 2282 TGAAGAGGCTGTGGCCACCTTATCCCTTTATGAAAAAAGAGAGAGAAACAGAG 2341  
Db 2613 TGAAGAGGCTGTGGCCACCTTATCCCTTTATGAAAAAAGAGAGAGAAACAGAG 2672  
Qy 2342 TGCTTAAACGACAGTGAAGAGAGGAGGACCTTTACAGGCGCACCATCTGCTGCGCACCTG 2401  
Db 2673 TGCTTAAACGACAGTGAAGAGAGGAGGACCTTTACAGGCGCACCATCTGCTGCGCACCTG 2732  
Qy 2402 TTAAGGCGGACGTCGACGACATAGGCAAGAAACATAGTTGGAGTGTCTTTGGTGTGAATA 2461  
Db 2733 TTAAGGCGGACGTCGACGACATAGGCAAGAAACATAGTTGGAGTGTCTTTGGTGTGAATA 2792  
Qy 2462 ATTTCCGAGTATTGATTTTAGGAGTCACTGCTCCATGATGAATGATCTGGAAGCTGCTC 2521  
Db 2793 ATTTCCGAGTATTGATTTTAGGAGTCACTGCTCCATGATGAATGATCTGGAAGCTGCTC 2852  
Qy 2522 TTGACCAACAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTCCCTGGATGAAA 2581  
Db 2853 TTGACCAACAAAGCAGATATAATTTGGCCCTGTGAGGACTCATCTCTCTCTCTGGATGAAA 2912

```
QY 2582 TGATTTTGTGTCACAGGAAATCGAGAGATTAGCTATAAGGATTCCATTGTTGATTGGAG 2641
Db 2913 TGATTTTGTGTCACAGGAAATCGAGAGATTAGCTATAAGGATTCCATTGTTGATTGGAG 2972
QY 2642 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTTG 2701
Db 2973 GAGCAACCACTTCAAAAACCCACACAGCAGTTAAAAATAGCTCCGAGATACAGTGCACTTG 3032
QY 2702 TAATCCATGTCCTGGACGCGTCAAGAGTGTGTGTGTTCCAGCTGTTAGATGAAA 2761
Db 3033 TAATCCATGTCCTGGACGCGTCAAGAGTGTGTGTGTTCCAGCTGTTAGATGAAA 3092
QY 2762 ATCTAAGGATGAATCTTTGAGGAATCATGGAAGATATGAAGATATTAAGACAGGCC 2821
Db 3093 ATCTAAGGATGAATCTTTGAGGAATCATGGAAGATATGAAGATATTAAGACAGGCC 3152
QY 2822 ATTATGAGTCTCTCAAGGAGAGAGATACCTTACCTTAACTCAAGCCAGAAAAAGTGGTT 2881
Db 3153 ATTATGAGTCTCTCAAGGAGAGAGATACCTTAACTCAAGCCAGAAAAANTNGTT 3212
QY 2882 TCCAAATGGATTGGCTGTGAACTCAACCAAGTGAAGCCACGTTTATTTGGGACCCAGG 2941
Db 3213 TCCAAATGGATTGGCTGTGAACTCAACCAAGTGAAGCCACGTTTATTTGGGACCCAGG 3272
QY 2942 TCTTTGAAGACTATGACCTGCAGAGCTGTGGAATGACTACATTTGACTGGAAGCCCTTCTTTG 3001
Db 3273 TCTTTGAAGACTATGACCTGCAGAGCTGTGGAATGACTACATTTGACTGGAAGCCCTTCTTTG 3332
QY 3002 ATGTCTGGCAGCTCCGGGGCAAGTACCCGGAATCGAGGCTTCCCAAGATATTTAAGACA 3061
Db 3333 ATGTCTGGCAGCTCCGGGGCAAGTACCCGGAATCGAGGCTTCCCAAGATATTTAAGACA 3392
QY 3062 AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3121
Db 3393 AAACAGTAGTGGAGAGCCAGGAAGTCTACGATGATGCCCAATATGCTGAACACAC 3452
QY 3122 TGATTAGTCAAAAGAACTCCGGCCCGGGGTGTGGTTGCTGGCCAGCACAGAGTA 3181
Db 3453 TGATTAGTCAAAAGAACTCCGGCCCGGGGTGTGGTTGCTGGCCAGCACAGAGTA 3512
QY 3182 TCCAAAGCAGCATTCACCTGTACAGAGGCTGTGTGCCCGCCAGGCTGCAGAGCCCATAG 3241
Db 3513 TCCAAAGCAGCATTCACCTGTACAGAGGCTGTGTGCCCGCCAGGCTGCAGAGCCCATAG 3572
QY 3242 CCACCTTTCTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3301
Db 3573 CCACCTTTCTATGGGTTAAGCAACAGGCTGAGAGGACTCTGCCAGCAGGAGCCATCT 3632
QY 3302 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTTGACTCTGGGCTGT 3361
Db 3633 ACTGCTCTCAGACTTCATCGCTCCCTTGCATTTCTGGCATCCGTTGACTCTGGGCTGT 3692
QY 3362 TTGCGCTTGCTGCTTTGGGTTAGAGAGCTGAGCAGGCTATGAGGATGATGGTGACG 3421
Db 3693 TTGCGCTTGCTGCTTTGGGTTAGAGAGCTGAGCAGGCTATGAGGATGATGGTGACG 3752
QY 3422 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGCCCTTTCAGAG 3481
Db 3753 ACTACAGCAGCATCATGCTCAAGCGCTGGGGGACCGGCTGGCAGAGCCCTTTCAGAG 3812
QY 3482 AGCTCCATGAAGAGTTGCCGAGNACTGTGGGCTTACTGTGGCAGTGAAGCCTGGACG 3541
Db 3813 AGCTCCATGAAGAGTTGCCGAGNACTGTGGGCTTACTGTGGCAGTGAAGCCTGGACG 3872
QY 3542 TCGCAGACCTCGGAAGTTGGGTTCAAGGCGATCCGCGCGGCTCCTGGCTACCCCAAGCC 3601
Db 3873 TCGCAGACCTCGGACGCTGGGTTCAAGGCGATCCGCGCGGCTCCTGGCTACCCCAAGCC 3932
QY 3602 AGCCCGACCAACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCAGCAGTCTACAG 3661
Db 3933 AGCCCGACCAACCGAGAGGCTCAACATGTGGAGACTCGCAGACATCAGCAGTCTACAG 3992
QY 3662 GCATTAGGTTAACAGAAATCATTAAGCAATGGGACCTGCTTCAGCAGTCTCAGGCTCTACT
```

```
Db 3993 GCATTAGGTTAACAGAAATCATTAAGCAATGGGACCTGCTTCAGCAGTCTCAGGCTCTACT 4052
QY 3722 TCTCCAAATTTGAAGTCCCAAAATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 3781
Db 4053 TCTCCAAATTTGAAGTCCCAAAATATTTGCTGTGGGGAAGATTTCCAGGATCAGGTTGAGG 4112
QY 3782 ATTATGCAATTGAGGGAAGAACATATCTGTGGCTGAGGTTGAGAAAATGGCTTGGACCCCATTT 3841
Db 4113 ATTATGCAATTGAGGGAAGAACATATCTGTGGCTGAGGTTGAGAAAATGGCTTGGACCCCATTT 4172
QY 3842 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 3875
Db 4173 TGGGATATGATACAGACTAACTTTTTTTTTTTTTT 4206
```

## RESULT 10

```
US-08-980-326-45/c
; Sequence 45, Application US/08980326
; Patent No. 6703197
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; FILE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER
; FILE REFERENCE: 50004/002003
; CURRENT APPLICATION NUMBER: US/08/980,326
; CURRENT FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 60/050,310
; EARLIER FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: 60/031,964
; EARLIER FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-980-326-45
```

```
Query Match 0.7%; Score 28; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2468 GAGTTATTGATTTAGGAGTCATGACTCC 2495
Db 28 GAGTTATTGATTTAGGAGTCATGACTCC 1
```

## RESULT 11

```
US-08-980-326-35
; Sequence 35, Application US/08980326
; Patent No. 6703197
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; FILE REFERENCE: 50004/002003
; CURRENT APPLICATION NUMBER: US/08/980,326
; CURRENT FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 60/050,310
; EARLIER FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: 60/031,964
; EARLIER FILING DATE: 1996-11-27
```

; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 35  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-35

Query Match 0.7%; Score 26; DB 3; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 TTCCGAGGTCAGGAATTAAAGATCA 239  
|||||  
Db 1 TTCCGAGGTCAGGAATTAAAGATCA 26

## RESULT 12

US-08-980-326-48  
; Sequence 48, Application US/08980326  
; Patent No. 6703197

; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.

; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel

; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric

; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,

; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003

; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20

; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27

; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 48  
; LENGTH: 26

; TYPE: DNA  
; ORGANISM: Homo sapiens

US-08-980-326-48

Query Match 0.7%; Score 26; DB 3; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2422 ATAGGCAAGAACATAGTTGGAGTAGT 2447  
|||||  
Db 1 ATAGGCAAGAACATAGTTGGAGTAGT 26

## RESULT 13

US-08-980-326-53

; Sequence 53, Application US/08980326  
; Patent No. 6703197

; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.

; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel

; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric

; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,

; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003

; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20

; EARLIER APPLICATION NUMBER: 60/031,964

; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 53  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-53

Query Match 0.7%; Score 26; DB 3; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2790 CATGGAAGAATATGAAGATATTAGAC 2815  
|||||  
Db 1 CATGGAAGAATATGAAGATATTAGAC 26

## RESULT 14

US-08-980-326-57/c

; Sequence 57, Application US/08980326  
; Patent No. 6703197

; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.

; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel

; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric

; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,

; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003

; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20

; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27

; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 57  
; LENGTH: 26

; TYPE: DNA  
; ORGANISM: Homo sapiens

US-08-980-326-57

Query Match 0.7%; Score 26; DB 3; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2496 ATGTGATAAGATACTGAAAGCTGCTC 2521  
|||||  
Db 26 ATGTGATAAGATACTGAAAGCTGCTC 1

## RESULT 15

US-08-980-326-63/c

; Sequence 63, Application US/08980326  
; Patent No. 6703197

; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.

; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel

; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric

; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,

; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003

; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20





; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 41  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-41

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2168 CCGGACCTCTCAATATAATTAAGG 2192  
Db 25 CCGGACCTCTCAATATAATTAAGG 1

RESULT 20  
US-08-980-326-49/c  
; Sequence 49, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 49  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-49

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2737 GTGTGTTCCAGCTGTTAGATGAAA 2761  
Db 25 GTGTGTTCCAGCTGTTAGATGAAA 1

RESULT 21  
US-08-980-326-58/c  
; Sequence 58, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310

; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-58

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1696 ACTGGAATGGAGGACACAACTTGT 1720  
Db 25 ACTGGAATGGAGGACACAACTTGT 1

RESULT 22  
US-08-980-326-59  
; Sequence 59, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 59  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-59

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1188 AGAGCGCTGTAATGTCAGGATCA 1212  
Db 1 AGAGCGCTGTAATGTCAGGATCA 25

RESULT 23  
US-08-980-326-60/c  
; Sequence 60, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26

/ EARLIER APPLICATION NUMBER: 60/050,310  
/ EARLIER FILING DATE: 1997-06-20  
/ EARLIER APPLICATION NUMBER: 60/031,964  
/ EARLIER FILING DATE: 1996-11-27  
/ NUMBER OF SEQ ID NOS: 75  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 60  
/ LENGTH: 25  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-08-980-326-60

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2096 CCCTGTGAAGGCGATTGAAACAA 2120  
|||||  
Db 25 CCCTGTGAAGGCGATTGAAACAA 1

## RESULT 24

US-08-980-326-61  
/ Sequence 61, Application US/08980326  
/ Patent No. 6703197  
/ GENERAL INFORMATION:  
/ APPLICANT: Gravel, Roy A.  
/ APPLICANT: Rozen, Rima  
/ APPLICANT: LeClerc, Daniel  
/ APPLICANT: Goyette, Philippe  
/ APPLICANT: Campeau, Eric  
/ TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
/ TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
/ TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
/ FILE REFERENCE: 50004/002003  
/ CURRENT APPLICATION NUMBER: US/08/980,326  
/ CURRENT FILING DATE: 1997-11-26  
/ EARLIER APPLICATION NUMBER: 60/050,310  
/ EARLIER FILING DATE: 1997-06-20  
/ EARLIER APPLICATION NUMBER: 60/031,964  
/ EARLIER FILING DATE: 1996-11-27  
/ NUMBER OF SEQ ID NOS: 75  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 61  
/ LENGTH: 25  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-08-980-326-61

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 TAAAGTATGAGCTGCTATGGTG 1551  
|||||  
Db 1 TAAAGTATGAGCTGCTATGGTG 25

## RESULT 25

US-08-980-326-62/c  
/ Sequence 62, Application US/08980326  
/ Patent No. 6703197  
/ GENERAL INFORMATION:  
/ APPLICANT: Gravel, Roy A.  
/ APPLICANT: Rozen, Rima  
/ APPLICANT: LeClerc, Daniel  
/ APPLICANT: Goyette, Philippe  
/ APPLICANT: Campeau, Eric  
/ TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
/ TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
/ TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
/ FILE REFERENCE: 50004/002003  
/ CURRENT APPLICATION NUMBER: US/08/980,326

/ CURRENT FILING DATE: 1997-11-26  
/ EARLIER APPLICATION NUMBER: 60/050,310  
/ EARLIER FILING DATE: 1997-06-20  
/ EARLIER APPLICATION NUMBER: 60/031,964  
/ EARLIER FILING DATE: 1996-11-27  
/ NUMBER OF SEQ ID NOS: 75  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 62  
/ LENGTH: 25  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-08-980-326-62

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1117 GAAGACATATGTACTGTCTGTC 1141  
|||||  
Db 25 GAAGACATATGTACTGTCTGTC 1

## RESULT 26

US-08-980-326-65/c  
/ Sequence 65, Application US/08980326  
/ Patent No. 6703197  
/ GENERAL INFORMATION:  
/ APPLICANT: Gravel, Roy A.  
/ APPLICANT: Rozen, Rima  
/ APPLICANT: LeClerc, Daniel  
/ APPLICANT: Goyette, Philippe  
/ APPLICANT: Campeau, Eric  
/ TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
/ TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
/ TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
/ FILE REFERENCE: 50004/002003  
/ CURRENT APPLICATION NUMBER: US/08/980,326  
/ CURRENT FILING DATE: 1997-11-26  
/ EARLIER APPLICATION NUMBER: 60/050,310  
/ EARLIER FILING DATE: 1997-06-20  
/ EARLIER APPLICATION NUMBER: 60/031,964  
/ EARLIER FILING DATE: 1996-11-27  
/ NUMBER OF SEQ ID NOS: 75  
/ SOFTWARE: FastSeq for Windows Version 3.0  
/ SEQ ID NO 65  
/ LENGTH: 25  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-08-980-326-65

Query Match 0.6%; Score 25; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3788 CATTGAGGAAGACATATCTGTGCG 3812  
|||||  
Db 25 CATTGAGGAAGACATATCTGTGCG 1

## RESULT 27

US-09-949-016-36633  
/ Sequence 36633, Application US/09949016  
/ Patent No. 6812339  
/ GENERAL INFORMATION:  
/ APPLICANT: VENTER, J. Craig et al.  
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
/ FILE REFERENCE: CL001307  
/ CURRENT APPLICATION NUMBER: US/09/949,016  
/ CURRENT FILING DATE: 2000-04-14  
/ PRIOR FILING DATE: 60/241,755  
/ PRIOR FILING DATE: 2000-10-20  
/ PRIOR APPLICATION NUMBER: 60/237,768

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36633
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36633

Query Match
Best Local Similarity 0.6%; Score 25; DB 3; Length 601;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3864 TTTTGTGCTTTTTCCTTTTGA 3888
Db 287 TTTTGTGCTTTTTCCTTTTGA 311

RESULT 28
US-09-949-016-44805
; Sequence 44805, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44805
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44805

Query Match
Best Local Similarity 0.6%; Score 25; DB 3; Length 601;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3864 TTTTGTGCTTTTTCCTTTTGA 3888
Db 287 TTTTGTGCTTTTTCCTTTTGA 311

RESULT 29
US-09-949-016-12696
; Sequence 12696, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 12696
; LENGTH: 15546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12696

Query Match
Best Local Similarity 0.6%; Score 25; DB 3; Length 15546;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3864 TTTTGTGCTTTTTCCTTTTGA 3888
Db 6839 TTTTGTGCTTTTTCCTTTTGA 6863

RESULT 30
US-09-949-016-13028
; Sequence 13028, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13028
; LENGTH: 15546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13028

Query Match
Best Local Similarity 0.6%; Score 25; DB 3; Length 15546;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3864 TTTTGTGCTTTTTCCTTTTGA 3888
Db 6839 TTTTGTGCTTTTTCCTTTTGA 6863

RESULT 31
US-09-949-016-16005/c
; Sequence 16005, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16005
; LENGTH: 24508
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16005
```

Query Match 0.6%; Score 25; DB 3; Length 24508;  
Best Local Similarity 100.0%; Pred. No. 0.94;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTGTGCTTTT 3887  
DB 1492 TTTTGTGCTTTTGTGCTTTT 1468

RESULT 32  
US-09-949-016-11864/c  
; Sequence 11864, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 11864  
; LENGTH: 24519  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11864

Query Match 0.6%; Score 25; DB 3; Length 24519;  
Best Local Similarity 100.0%; Pred. No. 0.94;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTGTGCTTTT 3887  
DB 1504 TTTTGTGCTTTTGTGCTTTT 1480

RESULT 33  
US-08-980-326-36/c  
; Sequence 36, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-36

Query Match 0.6%; Score 24; DB 3; Length 24;

Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCGGAGAGCTAAACGAGAACAC 213  
DB 24 CCGGAGAGCTAAACGAGAACAC 1

RESULT 34  
US-08-980-326-37/c  
; Sequence 37, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-37

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 ATAAGGAATCTTGCTGGCTGGG 331  
DB 24 ATAAGGAATCTTGCTGGCTGGG 1

RESULT 35  
US-08-980-326-38/c  
; Sequence 38, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 38  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-38

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3896 GATCCTCAAGGAATACACCTAG 3919  
Db 24 GATCCTCAAGGAATACACCTAG 1

## RESULT 36

US-08-980-326-40  
; Sequence 40, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 40  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-40

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2026 ACAGGAGGGAAGAACTCATTGAC 2049  
Db 1 ACAGGAGGGAAGAACTCATTGAC 24

## RESULT 37

US-08-980-326-42  
; Sequence 42, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 42  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-42

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CAACCCGAAGGTCTGAAGAAACC 114  
Db 1 CAACCCGAAGGTCTGAAGAAACC 24

## RESULT 38

US-08-980-326-50  
; Sequence 50, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 50  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-50

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3304 TGCCTCTCAGACTTCATCGCTCCC 3327  
Db 1 TGCCTCTCAGACTTCATCGCTCCC 24

## RESULT 39

US-08-980-326-52  
; Sequence 52, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 52  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-980-326-52

US-08-980-326-52

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2887 ATGATTGGCTGTCTGAACCTCAC 2910  
|||||  
DB 1 ATGATTGGCTGTCTGAACCTCAC 24

RESULT 40

US-08-980-326-54/c  
; Sequence 54, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 54  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-08-980-326-54

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3394 AGCAAGCCCTATGAGGATGATGTT 3417  
|||||  
DB 24 AGCAAGCCCTATGAGGATGATGTT 1

RESULT 41

US-08-980-326-56/c  
; Sequence 56, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 24  
; TYPE: DNA

; ORGANISM: Homo sapiens  
US-08-980-326-56

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2631 GTTGATTGGAGGAGCAACCACTTC 2654  
|||||  
DB 24 GTTGATTGGAGGAGCAACCACTTC 1

RESULT 42

US-08-980-326-64  
; Sequence 64, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 64  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-08-980-326-64

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3813 TGAAGTTGAGAAATGCGTTGGACC 3836  
|||||  
DB 1 TGAAGTTGAGAAATGCGTTGGACC 24

RESULT 43

US-08-980-326-66  
; Sequence 66, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 66  
; LENGTH: 24



```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-980-326-66

Query Match
Best Local Similarity 0.6%; Score 24; DB 3; Length 24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TGTGAGAGACCTCTCTCTGCC 31
Db 1 TGTGAGAGACCTCTCTCTGCC 24

RESULT 44
US-09-949-016-39482/c
; Sequence 39482, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39482
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-39482

Query Match
Best Local Similarity 0.6%; Score 24; DB 3; Length 601;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTGGCCTTTT 3886
Db 100 TTTTGTGCTTTTGGCCTTTT 77

RESULT 45
US-09-949-016-39515/c
; Sequence 39515, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39515
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-39515

Query Match
Best Local Similarity 0.6%; Score 24; DB 3; Length 601;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTGGCCTTTT 3886
Db 100 TTTTGTGCTTTTGGCCTTTT 77

RESULT 46
US-09-949-016-61884
; Sequence 61884, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61884
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61884

Query Match
Best Local Similarity 0.6%; Score 24; DB 3; Length 601;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3861 ACTTTTGTGCTTTTGGCCTTTT 3884
Db 493 ACTTTTGTGCTTTTGGCCTTTT 516

RESULT 47
US-09-949-016-63101/c
; Sequence 63101, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63101
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63101

Query Match
Best Local Similarity 0.6%; Score 24; DB 3; Length 601;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTGTGCTTTTGGCCTTTT 3886
Db 100 TTTTGTGCTTTTGGCCTTTT 77
```

```

; Sequence 118319, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118319
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-118319

Query Match          0.6%; Score 24; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. NO. 2.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTTTTTTTTTTTTGCCTTTT 3886
Db 100 TTTTTTTTTTTTTTTTGCCTTTT 77

Search completed: March 7, 2006, 02:21:28
Job time : 686 secs

```

```

Db      100 TTTTTCCTTTTTCCTTTT 77

RESULT 48
US-09-949-016-63123/c
; Sequence 63123, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63123
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63123

Query Match      0.6%; Score 24; DB 3; Length 601;
Best Local Similarity 100.0%; Pred.No. 2.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3863 TTTTTCCTTTTTCCTTTT 3886
          .|TTTTTTTTTTTTTTTTTTTT
Db      100 TTTTTCCTTTTTCCTTTT 77

```

```

RESULT 49
US-09-949-016-118281/c
; Sequence 118281, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118281
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-118281

Query Match          0.6%; Score 24; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3863 TTTTTTTTTTTTTTGGCTTTT 3886
Db      100 TTTTTTTTTTTTTTGGCTTTT 77

RESULT 50
US-09-949-016-118319/c

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 02:18:13 ; Search time 2700 Seconds  
(without alignments)  
12002.849 Million cell updates/sec

Title: US-10-607-712-1  
Perfect score: 3919  
Sequence: 1 ggtcaactgtgagacg.....ctcaaggaaatacaactag 3919

Scoring table:

OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 18

Total number of hits satisfying chosen parameters: 57335

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA Main.\*

- 1: /cgm2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgm2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 3: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /cgm2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /cgm2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 6: /cgm2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 7: /cgm2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 8: /cgm2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 9: /cgm2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 10: /cgm2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3919	100.0	3919	7	US-10-607-712-1
2	3601	91.9	3856	7	US-10-607-712-75
3	318	8.1	420	3	US-09-796-682-8171
4	318	8.1	420	5	US-10-040-862-8171
5	318	8.1	420	6	US-10-057-475B-8171
6	318	8.1	420	6	US-10-154-884B-8171
7	318	8.1	420	8	US-10-764-324-8171
8	222	5.7	305	5	US-10-040-739-642
9	124	3.2	901	5	US-10-027-632-121776
10	124	3.2	901	6	US-10-027-632-121776
11	124	3.2	1275	4	US-09-925-065A-727226
12	80	2.0	616	4	US-09-925-065A-697175
13	69	1.8	493	5	US-10-027-632-56722
14	69	1.8	493	5	US-10-027-632-314464
15	69	1.8	493	5	US-10-027-632-314465
16	69	1.8	493	6	US-10-027-632-314464
17	69	1.8	493	6	US-10-027-632-314465
18	69	1.8	627	5	US-10-027-632-3760
19	69	1.8	627	5	US-10-027-632-3760
20	69	1.8	627	5	US-10-027-632-3760
21	39	1.0	611	4	US-09-925-065A-460077
22	31	0.8	588	4	US-09-925-065A-39453
23	28	0.7	28	7	US-10-607-712-45

C 97	24	0.6	479	5	US-10-027-632-310427	Sequence 310427,	170	23	0.6	717	5	US-10-027-632-107263	Sequence 107263,
C 98	24	0.6	479	6	US-10-027-632-57379	Sequence 57379, A	171	23	0.6	717	6	US-10-027-632-107263	Sequence 107262,
C 99	24	0.6	479	6	US-10-027-632-310427	Sequence 310427,	172	23	0.6	717	6	US-10-027-632-107263	Sequence 107263,
C 100	24	0.6	489	4	US-09-925-065A-412268	Sequence 412268,	C 173	23	0.6	821	5	US-10-027-632-153100	Sequence 153100,
C 101	24	0.6	506	4	US-09-925-065A-270554	Sequence 270554,	C 174	23	0.6	821	5	US-10-027-632-153101	Sequence 153101,
C 102	24	0.6	512	4	US-09-925-065A-620176	Sequence 620176,	C 175	23	0.6	821	6	US-10-027-632-153100	Sequence 153100,
C 103	24	0.6	548	3	US-09-963-034-2933	Sequence 2933, Ap	C 176	23	0.6	821	6	US-10-027-632-153101	Sequence 153101,
C 104	24	0.6	548	4	US-09-925-065A-424567	Sequence 424567,	C 177	23	0.6	825	8	US-10-425-115-85657	Sequence 85657, A
C 105	24	0.6	549	4	US-09-925-065A-889241	Sequence 889241,	C 178	23	0.6	885	8	US-10-425-115-76054	Sequence 76054, A
C 106	24	0.6	585	4	US-09-925-065A-624935	Sequence 624935,	C 179	23	0.6	955	4	US-09-925-065A-72630	Sequence 72630, A
C 107	24	0.6	626	4	US-09-925-065A-530520	Sequence 530520,	C 180	23	0.6	1395	4	US-09-925-065A-78441	Sequence 78441, A
C 108	24	0.6	655	5	US-10-027-632-27407	Sequence 27407, A	C 181	23	0.6	1610	6	US-10-133-937-77	Sequence 77, Appl
C 109	24	0.6	655	5	US-10-027-632-27408	Sequence 27408, A	C 182	23	0.6	1610	6	US-10-159-563-77	Sequence 77, Appl
C 110	24	0.6	655	6	US-10-027-632-27407	Sequence 27407, A	C 183	23	0.6	1875	5	US-10-106-698-222	Sequence 222, App
C 111	24	0.6	655	6	US-10-027-632-27408	Sequence 27408, A	C 184	23	0.6	2073	5	US-10-027-632-97481	Sequence 97481, A
C 112	24	0.6	766	5	US-10-027-632-5085	Sequence 5085, Ap	C 185	23	0.6	2073	6	US-10-027-632-97481	Sequence 97481, A
C 113	24	0.6	786	6	US-10-027-632-5085	Sequence 5085, Ap	C 186	23	0.6	3382	6	US-10-443-201-27	Sequence 27, Appl
C 114	24	0.6	794	7	US-10-437-963-80007	Sequence 80007, A	C 187	23	0.6	4840	8	US-10-723-860-6164	Sequence 6164, Ap
C 115	24	0.6	890	5	US-10-011-585A-134	Sequence 134, App	C 188	23	0.6	5565	6	US-10-311-455-236	Sequence 236, App
C 116	24	0.6	9417	9	US-10-853-335A-22	Sequence 22, Appl	C 189	23	0.6	9456	8	US-10-723-860-6147	Sequence 6147, Ap
C 117	24	0.6	32249	6	US-10-242-355-1132	Sequence 1132, Ap	C 190	23	0.6	10480	6	US-10-311-455-2174	Sequence 2174, Ap
C 118	24	0.6	86500	8	US-10-713-933-6771	Sequence 6771, Ap	C 191	23	0.6	14117	3	US-09-954-531-1377	Sequence 1377, Ap
C 119	24	0.6	137908	9	US-10-287-436A-779	Sequence 779, App	C 192	23	0.6	14117	9	US-10-843-641A-2444	Sequence 2444, Ap
C 120	24	0.6	193853	5	US-10-087-192-1663	Sequence 1663, Ap	C 193	23	0.6	17431	3	US-09-764-855-247	Sequence 247, App
C 121	24	0.6	241805	7	US-10-741-601-5621	Sequence 5621, Ap	C 194	23	0.6	17431	5	US-10-072-349-247	Sequence 247, App
C 122	24	0.6	241805	8	US-10-741-600-17581	Sequence 17581, A	C 195	23	0.6	20486	6	US-10-240-485-164	Sequence 164, App
C 123	24	0.6	325446	8	US-10-719-993-6824	Sequence 6824, Ap	C 196	23	0.6	34245	7	US-10-367-094-102	Sequence 102, App
C 124	23	0.6	23	7	US-10-607-712-46	Sequence 46, Appl	C 197	23	0.6	81924	7	US-10-741-601-5636	Sequence 5636, Ap
C 125	23	0.6	23	7	US-10-607-712-55	Sequence 55, Appl	C 198	23	0.6	107910	9	US-10-981-277-19	Sequence 19, Appl
C 126	23	0.6	201	7	US-10-741-601-14799	Sequence 14799, A	C 199	23	0.6	169659	7	US-10-322-696-70	Sequence 70, Appl
C 127	23	0.6	201	8	US-10-741-600-25758	Sequence 25758, A	C 200	23	0.6	203127	7	US-10-741-601-5654	Sequence 5654, Ap
C 128	23	0.6	201	8	US-10-741-600-25759	Sequence 25759, A	C 201	23	0.6	203127	8	US-10-741-600-17666	Sequence 17666, A
C 129	23	0.6	201	8	US-10-741-600-25764	Sequence 25764, A	C 202	23	0.6	510510	8	US-10-741-600-17606	Sequence 17606, A
C 130	23	0.6	201	8	US-10-741-600-37188	Sequence 37188, A	C 203	22	0.6	22	7	US-10-607-712-47	Sequence 47, Appl
C 131	23	0.6	225	9	US-10-933-118-3	Sequence 3, Appl	C 204	22	0.6	194	3	US-09-867-701-9917	Sequence 9917, Ap
C 132	23	0.6	270	8	US-10-425-115-158083	Sequence 158083,	C 205	22	0.6	201	8	US-10-719-993-30838	Sequence 30838, A
C 133	23	0.6	282	8	US-10-425-115-92332	Sequence 92332, A	C 206	22	0.6	201	8	US-10-719-993-31678	Sequence 31678, A
C 134	23	0.6	361	8	US-10-425-115-88628	Sequence 88628, A	C 207	22	0.6	244	8	US-10-425-115-98422	Sequence 98422, A
C 135	23	0.6	376	4	US-09-925-065A-261182	Sequence 261182,	C 208	22	0.6	250	7	US-10-242-535A-54640	Sequence 54640, A
C 136	23	0.6	385	7	US-10-152-319A-57	Sequence 4517, App	C 209	22	0.6	250	7	US-10-242-535A-54648	Sequence 54648, A
C 137	23	0.6	421	3	US-09-917-800A-831	Sequence 831, App	C 210	22	0.6	250	7	US-10-085-783A-54648	Sequence 54648, A
C 138	23	0.6	421	6	US-10-388-934-31	Sequence 31, Appl	C 211	22	0.6	250	7	US-10-085-783A-54648	Sequence 54648, A
C 139	23	0.6	434	3	US-09-876-143-174	Sequence 174, App	C 212	22	0.6	287	3	US-09-867-701-8939	Sequence 8939, Ap
C 140	23	0.6	434	3	US-09-876-143-262	Sequence 262, App	C 213	22	0.6	324	3	US-09-917-800A-641	Sequence 641, App
C 141	23	0.6	451	4	US-09-925-065A-351757	Sequence 351757,	C 214	22	0.6	339	7	US-10-437-963-34514	Sequence 34514, A
C 142	23	0.6	459	7	US-10-424-599-116134	Sequence 116134,	C 215	22	0.6	350	7	US-10-437-963-34514	Sequence 34514, A
C 143	23	0.6	460	8	US-10-357-930-56449	Sequence 56449, A	C 216	22	0.6	353	7	US-10-424-599-84827	Sequence 84827, A
C 144	23	0.6	481	3	US-09-746-783-18	Sequence 18, Appl	C 217	22	0.6	374	6	US-10-424-599-80394	Sequence 80394, A
C 145	23	0.6	498	4	US-09-925-065A-98281	Sequence 98281, A	C 218	22	0.6	374	8	US-10-058-053A-4	Sequence 4, Appl
C 146	23	0.6	498	4	US-09-925-065A-98282	Sequence 98282, A	C 219	22	0.6	424	4	US-09-925-065A-204738	Sequence 204738,
C 147	23	0.6	503	4	US-09-925-065A-925330	Sequence 925330,	C 220	22	0.6	424	4	US-09-925-065A-204739	Sequence 204739,
C 148	23	0.6	522	4	US-09-925-065A-828324	Sequence 828324,	C 221	22	0.6	424	8	US-10-357-930-13696	Sequence 13696, A
C 149	23	0.6	523	4	US-09-925-065A-819272	Sequence 819272,	C 222	22	0.6	436	8	US-10-357-930-34836	Sequence 34836, A
C 150	23	0.6	529	8	US-10-425-115-105702	Sequence 105702,	C 223	22	0.6	452	7	US-10-424-599-67712	Sequence 67712, A
C 151	23	0.6	556	4	US-09-925-065A-3410	Sequence 3410, Ap	C 224	22	0.6	468	5	US-10-027-632-201664	Sequence 201664,
C 152	23	0.6	574	4	US-09-925-065A-755600	Sequence 755600,	C 225	22	0.6	468	5	US-10-027-632-201664	Sequence 201664,
C 153	23	0.6	575	4	US-09-925-065A-278724	Sequence 278724,	C 226	22	0.6	481	4	US-09-925-065A-401865	Sequence 401865,
C 154	23	0.6	576	4	US-09-925-065A-561677	Sequence 561677,	C 227	22	0.6	488	3	US-09-918-995-24463	Sequence 24463, A
C 155	23	0.6	600	9	US-10-972-079-24722	Sequence 24722, A	C 228	22	0.6	493	3	US-09-918-995-21166	Sequence 21166, A
C 156	23	0.6	600	9	US-10-972-079-24723	Sequence 24723, A	C 229	22	0.6	494	4	US-09-925-065A-396325	Sequence 396325,
C 157	23	0.6	600	9	US-10-972-079-73651	Sequence 73651, A	C 230	22	0.6	494	4	US-09-925-065A-396326	Sequence 396326,
C 158	23	0.6	600	9	US-10-972-079-73652	Sequence 73652, A	C 231	22	0.6	494	5	US-10-027-632-231843	Sequence 231843,
C 159	23	0.6	601	4	US-09-925-065A-478484	Sequence 478484,	C 232	22	0.6	494	6	US-10-027-632-231843	Sequence 231843,
C 160	23	0.6	601	4	US-09-925-065A-478485	Sequence 478485,	C 233	22	0.6	505	4	US-09-925-065A-77965	Sequence 77965, A
C 161	23	0.6	611	4	US-09-925-065A-188252	Sequence 188252,	C 234	22	0.6	506	3	US-09-814-353-13630	Sequence 13630, A
C 162	23	0.6	613	4	US-09-925-065A-188253	Sequence 188253,	C 235	22	0.6	514	4	US-09-925-065A-113707	Sequence 113707,
C 163	23	0.6	619	5	US-10-925-065A-751693	Sequence 751693,	C 236	22	0.6	518	3	US-09-873-367C-1045	Sequence 1045, Ap
C 164	23	0.6	619	6	US-10-027-632-273790	Sequence 273790,	C 237	22	0.6	518	9	US-10-843-641A-1045	Sequence 1045, Ap
C 165	23	0.6	634	4	US-09-925-065A-739010	Sequence 739010,	C 238	22	0.6	519	4	US-09-925-065A-417525	Sequence 417525,
C 166	23	0.6	705	4	US-09-925-065A-922916	Sequence 922916,	C 239	22	0.6	519	4	US-09-925-065A-417527	Sequence 417527,
C 167	23	0.6	716	5	US-10-027-632-159784	Sequence 159784,	C 240	22	0.6	520	4	US-09-925-065A-42413	Sequence 42413, A
C 168	23	0.6	716	6	US-10-027-632-159784	Sequence 159784,	C 241	22	0.6	524	4	US-09-925-065A-583132	Sequence 583132,
C 169	23	0.6	717	5	US-10-027-632-107262	Sequence 107262,	C 242	22	0.6	524	4	US-09-925-065A-650278	Sequence 650278,

C 243	22	0.6	531	5	US-10-027-632-210314	Sequence 210314,	C 316	22	0.6	982	4	US-09-925-065A-725590	Sequence 725590,
C 244	22	0.6	531	6	US-10-027-632-210314	Sequence 210314,	C 317	22	0.6	998	8	US-10-425-115-68031	Sequence 68031, A
C 245	22	0.6	535	4	US-09-925-065A-417523	Sequence 417523,	C 318	22	0.6	1031	9	US-10-956-157-17660	Sequence 17660, Ap
C 246	22	0.6	535	4	US-09-925-065A-417524	Sequence 417524,	C 319	22	0.6	1031	9	US-10-956-157-6995	Sequence 6995, Ap
C 247	22	0.6	536	4	US-09-925-065A-924158	Sequence 924158,	C 320	22	0.6	1136	6	US-10-320-352-1	Sequence 1, Appli
C 248	22	0.6	538	4	US-09-925-065A-102798	Sequence 102798,	C 321	22	0.6	1170	4	US-09-925-065A-297981	Sequence 297981,
C 249	22	0.6	538	4	US-09-925-065A-102799	Sequence 102799,	C 322	22	0.6	1225	4	US-09-925-065A-703263	Sequence 703263,
C 250	22	0.6	540	4	US-09-925-065A-11434	Sequence 11434, A	C 323	22	0.6	1228	5	US-10-027-632-253407	Sequence 253407,
C 251	22	0.6	540	4	US-09-925-065A-105964	Sequence 105964,	C 324	22	0.6	1228	6	US-10-027-632-253407	Sequence 253407,
C 252	22	0.6	550	4	US-09-925-065A-262968	Sequence 262968,	C 325	22	0.6	1352	9	US-10-314-942-9	Sequence 9, Appli
C 253	22	0.6	560	7	US-10-152-319A-825	Sequence 825, App	C 326	22	0.6	1470	8	US-10-775-920-155	Sequence 155, App
C 254	22	0.6	563	4	US-09-925-065A-89853	Sequence 89853,	C 327	22	0.6	1471	9	US-10-504-173-9	Sequence 9, Appli
C 255	22	0.6	569	4	US-09-925-065A-112247	Sequence 112247,	C 328	22	0.6	1518	3	US-09-729-674-143	Sequence 143, App
C 256	22	0.6	569	4	US-09-925-065A-112248	Sequence 112248,	C 329	22	0.6	1518	3	US-10-913-553-143	Sequence 143, App
C 257	22	0.6	569	6	US-10-101-510-99	Sequence 99, Appl	C 330	22	0.6	1520	3	US-09-397-945-101	Sequence 101, App
C 258	22	0.6	569	7	US-10-283-975A-32	Sequence 32,	C 331	22	0.6	1520	7	US-10-653-595-101	Sequence 101, App
C 259	22	0.6	570	4	US-09-925-065A-314680	Sequence 314680,	C 332	22	0.6	1573	3	US-09-746-783-187	Sequence 187, App
C 260	22	0.6	570	4	US-09-925-065A-314681	Sequence 314681,	C 333	22	0.6	1739	5	US-10-236-433-6	Sequence 6, Appli
C 261	22	0.6	573	7	US-10-437-963-40312	Sequence 40312, A	C 334	22	0.6	1766	4	US-09-925-065A-30791	Sequence 30791, A
C 262	22	0.6	578	4	US-09-925-065A-297980	Sequence 297980,	C 335	22	0.6	1791	7	US-10-425-114-32464	Sequence 32464, A
C 263	22	0.6	585	8	US-10-425-115-120425	Sequence 120425,	C 336	22	0.6	1808	8	US-10-425-115-59855	Sequence 59855, A
C 264	22	0.6	588	3	US-09-871-161-275	Sequence 275, App	C 337	22	0.6	1809	8	US-10-723-860-5723	Sequence 5723, Ap
C 265	22	0.6	589	4	US-09-925-065A-878937	Sequence 878937,	C 338	22	0.6	1809	8	US-10-723-860-6360	Sequence 6360, Ap
C 266	22	0.6	589	4	US-09-925-065A-539793	Sequence 539793,	C 339	22	0.6	1999	3	US-09-873-367C-1065	Sequence 1065, Ap
C 267	22	0.6	598	8	US-10-357-930-56158	Sequence 56158, A	C 340	22	0.6	1999	9	US-10-843-641A-1065	Sequence 1065, Ap
C 268	22	0.6	599	9	US-10-972-079-18874	Sequence 18874, A	C 341	22	0.6	2274	10	US-11-097-143-33417	Sequence 33417, A
C 269	22	0.6	600	4	US-09-925-065A-902706	Sequence 902706,	C 342	22	0.6	2315	4	US-09-925-065A-706810	Sequence 706810,
C 270	22	0.6	601	4	US-09-925-065A-575787	Sequence 575787,	C 343	22	0.6	2315	4	US-09-925-065A-706811	Sequence 706811,
C 271	22	0.6	605	4	US-09-925-065A-898473	Sequence 898473,	C 344	22	0.6	2351	6	US-10-108-260A-1701	Sequence 1701, Ap
C 272	22	0.6	608	4	US-09-925-065A-637147	Sequence 637147,	C 345	22	0.6	2595	5	US-10-106-698-451	Sequence 451, App
C 273	22	0.6	609	4	US-09-925-065A-809765	Sequence 809765,	C 346	22	0.6	2628	4	US-09-925-065A-676638	Sequence 676638,
C 274	22	0.6	612	5	US-10-027-632-71623	Sequence 71623, A	C 347	22	0.6	3002	4	US-09-925-065A-674422	Sequence 674422,
C 275	22	0.6	612	5	US-10-027-632-71623	Sequence 71623, A	C 348	22	0.6	3002	4	US-09-925-065A-674423	Sequence 674423,
C 276	22	0.6	614	5	US-10-027-632-209251	Sequence 209251,	C 349	22	0.6	3115	3	US-09-925-299-123	Sequence 123, App
C 277	22	0.6	614	5	US-10-027-632-209252	Sequence 209252,	C 350	22	0.6	3115	7	US-09-925-299-123	Sequence 123, App
C 278	22	0.6	614	6	US-10-027-632-209251	Sequence 209251,	C 351	22	0.6	3157	7	US-10-282-122A-36766	Sequence 36766, A
C 279	22	0.6	614	6	US-10-027-632-209251	Sequence 209252,	C 352	22	0.6	3255	5	US-10-027-632-115124	Sequence 115124,
C 280	22	0.6	619	7	US-10-242-535A-13956	Sequence 13956, A	C 353	22	0.6	3255	6	US-10-027-632-115124	Sequence 115124,
C 281	22	0.6	619	7	US-10-085-783A-13956	Sequence 13956, A	C 354	22	0.6	3452	8	US-10-723-860-7644	Sequence 7644, Ap
C 282	22	0.6	635	4	US-09-925-065A-731865	Sequence 731865,	C 355	22	0.6	3788	8	US-10-425-115-45050	Sequence 45050, A
C 283	22	0.6	655	5	US-10-027-632-6052	Sequence 6052, Ap	C 356	22	0.6	4182	3	US-09-814-353-20319	Sequence 20319, A
C 284	22	0.6	655	5	US-10-027-632-6053	Sequence 6053, Ap	C 357	22	0.6	4256	10	US-11-097-143-34252	Sequence 34252, A
C 285	22	0.6	655	6	US-10-027-632-6052	Sequence 6052, Ap	C 358	22	0.6	5413	7	US-10-221-714A-418	Sequence 418, App
C 286	22	0.6	655	6	US-10-027-632-6053	Sequence 6053, Ap	C 359	22	0.6	6450	5	US-10-096-710-2	Sequence 2, Appli
C 287	22	0.6	657	7	US-10-424-599-9950	Sequence 9950, Ap	C 360	22	0.6	6450	5	US-10-081-563-1	Sequence 1, Appli
C 288	22	0.6	659	8	US-10-425-115-148511	Sequence 148511,	C 361	22	0.6	6450	5	US-10-052-092-1	Sequence 1, Appli
C 289	22	0.6	666	5	US-10-027-632-110414	Sequence 110414,	C 362	22	0.6	6450	5	US-10-052-092-7	Sequence 7, Appli
C 290	22	0.6	666	5	US-10-027-632-110414	Sequence 110414,	C 363	22	0.6	6450	5	US-10-207-655-60	Sequence 60, Appli
C 291	22	0.6	688	5	US-10-066-543-1210	Sequence 1210, Ap	C 364	22	0.6	6450	6	US-10-177-293-127	Sequence 127, App
C 292	22	0.6	714	7	US-10-424-599-65277	Sequence 65277, A	C 365	22	0.6	6450	6	US-10-027-983-3	Sequence 3, Appli
C 293	22	0.6	734	4	US-09-925-065A-80511	Sequence 80511, A	C 366	22	0.6	6450	6	US-10-007-926A-422	Sequence 422, App
C 294	22	0.6	734	4	US-09-925-065A-80512	Sequence 80512, A	C 367	22	0.6	6450	6	US-10-437-107-1	Sequence 1, Appli
C 295	22	0.6	759	3	US-09-919-580-825	Sequence 825, App	C 368	22	0.6	6450	6	US-10-437-107-7	Sequence 7, Appli
C 296	22	0.6	795	5	US-10-027-632-163308	Sequence 163308,	C 369	22	0.6	6450	6	US-10-392-274-1	Sequence 1, Appli
C 297	22	0.6	795	5	US-10-027-632-163309	Sequence 163309,	C 370	22	0.6	6450	6	US-10-448-753-3	Sequence 3, Appli
C 298	22	0.6	795	5	US-10-027-632-163310	Sequence 163310,	C 371	22	0.6	6450	6	US-10-172-118-435	Sequence 435, App
C 299	22	0.6	795	6	US-10-027-632-163308	Sequence 163308,	C 372	22	0.6	6450	6	US-10-388-360-303	Sequence 303, App
C 300	22	0.6	795	6	US-10-027-632-163309	Sequence 163309,	C 373	22	0.6	6450	7	US-10-342-887-435	Sequence 435, App
C 301	22	0.6	795	6	US-10-027-632-163310	Sequence 163310,	C 374	22	0.6	6450	8	US-10-370-159-75	Sequence 75, Appli
C 302	22	0.6	796	5	US-10-027-632-142106	Sequence 142106,	C 375	22	0.6	6450	8	US-10-896-419-1	Sequence 1, Appli
C 303	22	0.6	796	5	US-10-027-632-165236	Sequence 165236,	C 376	22	0.6	6450	8	US-10-896-419-7	Sequence 7, Appli
C 304	22	0.6	796	6	US-10-027-632-165237	Sequence 165237,	C 377	22	0.6	6450	9	US-10-928-465-72	Sequence 72, Appli
C 305	22	0.6	796	6	US-10-027-632-142106	Sequence 142106,	C 378	22	0.6	6450	9	US-10-888-313A-100	Sequence 100, App
C 306	22	0.6	796	6	US-10-027-632-165236	Sequence 165236,	C 379	22	0.6	6619	5	US-10-723-860-6901	Sequence 6901, Ap
C 307	22	0.6	796	6	US-10-027-632-165237	Sequence 165237,	C 380	22	0.6	7510	5	US-10-001-835-113	Sequence 113, App
C 308	22	0.6	810	8	US-10-425-115-71750	Sequence 71750, A	C 381	22	0.6	8042	6	US-10-311-455-900	Sequence 900, App
C 309	22	0.6	816	7	US-10-152-319A-1717	Sequence 1717, Ap	C 382	22	0.6	8168	6	US-10-311-455-901	Sequence 901, App
C 310	22	0.6	844	7	US-10-424-599-25219	Sequence 25219, A	C 383	22	0.6	8168	7	US-10-240-454-23	Sequence 23, Appli
C 311	22	0.6	876	7	US-10-425-114-30264	Sequence 30264, A	C 384	22	0.6	10957	6	US-10-311-455-1084	Sequence 1084, Ap
C 312	22	0.6	905	7	US-10-424-599-78300	Sequence 78300, A	C 385	22	0.6	11036	5	US-10-239-676-118	Sequence 118, App
C 313	22	0.6	905	7	US-10-437-963-80008	Sequence 80008, A	C 386	22	0.6	11036	6	US-10-240-453-138	Sequence 138, App
C 314	22	0.6	952	4	US-09-925-065A-701979	Sequence 701979,	C 387	22	0.6	12405	5	US-10-239-676-35	Sequence 35, Appli
C 315	22	0.6	980	4	US-09-925-065A-696097	Sequence 696097,	C 388	22	0.6	12405	6	US-10-240-453-43	Sequence 43, Appli

C 389	22	0.6	12405	7	US-10-221-613-101	Sequence 101, App	C 462	21	0.5	201	7	US-10-741-601-6881	Sequence 6881, Ap
C 390	22	0.6	13182	3	US-09-817-199A-3	Sequence 3, Appli	463	21	0.5	201	8	US-10-719-993-18653	Sequence 18653, A
C 391	22	0.6	13182	10	US-11-043-964-3	Sequence 3, Appli	464	21	0.5	201	8	US-10-719-993-19852	Sequence 19852, A
C 392	22	0.6	17142	5	US-10-239-676-206	Sequence 206, App	465	21	0.5	201	8	US-10-719-993-19853	Sequence 19853, A
C 393	22	0.6	17142	6	US-10-311-455-2080	Sequence 2080, Ap	466	21	0.5	201	8	US-10-719-993-19854	Sequence 19854, A
C 394	22	0.6	17142	6	US-10-240-453-304	Sequence 304, App	C 467	21	0.5	201	8	US-10-719-993-21639	Sequence 21639, A
C 395	22	0.6	17200	3	US-09-764-877-3390	Sequence 3390, Ap	468	21	0.5	201	8	US-10-719-993-25064	Sequence 25064, A
C 396	22	0.6	17200	3	US-10-242-515-3390	Sequence 3390, Ap	469	21	0.5	201	8	US-10-719-993-43378	Sequence 43378, A
C 397	22	0.6	17528	6	US-10-311-455-574	Sequence 574, App	470	21	0.5	201	8	US-10-719-993-45753	Sequence 45753, A
C 398	22	0.6	19338	9	US-10-737-082-99	Sequence 99, Appli	471	21	0.5	201	8	US-10-719-993-53040	Sequence 53040, A
C 399	22	0.6	19338	9	US-10-765-790-99	Sequence 99, Appli	472	21	0.5	201	8	US-10-719-993-53760	Sequence 53760, A
C 400	22	0.6	22478	3	US-09-997-722-115	Sequence 115, App	473	21	0.5	201	8	US-10-719-993-53762	Sequence 53762, A
C 401	22	0.6	22609	3	US-09-764-877-2146	Sequence 2146, Ap	474	21	0.5	201	8	US-10-719-993-53764	Sequence 53764, A
C 402	22	0.6	22609	6	US-10-242-515-2146	Sequence 2146, Ap	475	21	0.5	201	8	US-10-719-993-54312	Sequence 54312, A
C 403	22	0.6	24707	3	US-09-740-027-3	Sequence 3, Appli	C 476	21	0.5	201	8	US-10-741-600-19377	Sequence 19377, A
C 404	22	0.6	24707	5	US-10-274-968-3	Sequence 3, Appli	C 477	21	0.5	201	8	US-10-741-600-19378	Sequence 19378, A
C 405	22	0.6	24707	10	US-11-043-932-3	Sequence 3, Appli	478	21	0.5	201	8	US-10-741-600-38488	Sequence 38488, A
C 406	22	0.6	32193	3	US-09-764-877-2147	Sequence 2147, Ap	479	21	0.5	201	8	US-10-741-600-39687	Sequence 39687, A
C 407	22	0.6	32193	6	US-10-242-515-2147	Sequence 2147, Ap	480	21	0.5	201	8	US-10-741-600-39688	Sequence 39688, A
C 408	22	0.6	35000	3	US-09-793-807-10	Sequence 10, Appli	481	21	0.5	201	8	US-10-741-600-39689	Sequence 39689, A
C 409	22	0.6	36445	10	US-11-097-143-24415	Sequence 24415, A	C 482	21	0.5	201	8	US-10-741-600-41472	Sequence 41472, A
C 410	22	0.6	38054	5	US-10-087-192-1720	Sequence 1720, Ap	483	21	0.5	201	8	US-10-741-600-65895	Sequence 65895, A
C 411	22	0.6	41787	7	US-10-087-192-1720	Sequence 1720, Ap	484	21	0.5	201	8	US-10-741-600-71170	Sequence 71170, A
C 412	22	0.6	42334	7	US-10-322-281-820	Sequence 820, App	485	21	0.5	201	8	US-10-741-600-71172	Sequence 71172, A
C 413	22	0.6	42334	7	US-10-052-482-94	Sequence 94, Appli	486	21	0.5	201	8	US-10-741-600-71174	Sequence 71174, A
C 414	22	0.6	43887	7	US-10-322-281-109	Sequence 109, App	487	21	0.5	201	8	US-10-741-600-71722	Sequence 71722, A
C 415	22	0.6	46148	8	US-10-719-993-7064	Sequence 7064, Ap	C 488	21	0.5	220	7	US-10-242-535A-2201	Sequence 2201, Ap
C 416	22	0.6	54016	8	US-10-741-600-17886	Sequence 17886, A	C 489	21	0.5	220	7	US-10-085-783A-2201	Sequence 2201, Ap
C 417	22	0.6	59475	7	US-10-322-696-166	Sequence 166, App	490	21	0.5	248	8	US-10-425-115-10205	Sequence 10205, A
C 418	22	0.6	61197	5	US-10-087-192-1924	Sequence 1924, Ap	491	21	0.5	248	8	US-10-425-115-140881	Sequence 140881, A
C 419	22	0.6	75252	5	US-10-087-192-904	Sequence 904, App	C 492	21	0.5	285	4	US-09-925-065A-130263	Sequence 130263, A
C 420	22	0.6	76180	7	US-10-322-281-492	Sequence 492, App	C 493	21	0.5	290	8	US-10-357-930-19339	Sequence 19339, A
C 421	22	0.6	79528	8	US-10-723-860-2621	Sequence 2621, Ap	494	21	0.5	292	8	US-10-357-930-7817	Sequence 7817, Ap
C 422	22	0.6	79528	9	US-10-756-149-2427	Sequence 2427, Ap	C 495	21	0.5	296	7	US-10-242-535A-52499	Sequence 52499, A
C 423	22	0.6	81210	8	US-10-741-601-5664	Sequence 5664, Ap	C 496	21	0.5	296	7	US-10-085-783A-52499	Sequence 52499, A
C 424	22	0.6	87467	7	US-10-741-601-5634	Sequence 17691, A	C 497	21	0.5	300	5	US-10-079-623-311	Sequence 311, App
C 425	22	0.6	87467	7	US-10-741-601-5634	Sequence 5634, Ap	498	21	0.5	310	7	US-10-191-803-856	Sequence 856, App
C 426	22	0.6	87467	8	US-10-741-600-17624	Sequence 17624, A	C 499	21	0.5	315	3	US-09-960-352-12366	Sequence 12366, A
C 427	22	0.6	105413	6	US-10-427-923-3	Sequence 3, Appli	500	21	0.5	319	8	US-10-674-124A-4819	Sequence 4819, Ap
C 428	22	0.6	105413	10	US-11-124-082-3	Sequence 3, Appli	501	21	0.5	327	3	US-09-960-352-14937	Sequence 14937, A
C 429	22	0.6	110096	3	US-09-880-107-1542	Sequence 1542, Ap	C 502	21	0.5	338	5	US-10-079-623-131	Sequence 131, App
C 430	22	0.6	150319	9	US-10-981-277-37	Sequence 37, Appli	C 503	21	0.5	339	8	US-10-357-930-19637	Sequence 19627, A
C 431	22	0.6	156416	9	US-10-461-862-12	Sequence 12, Appli	C 504	21	0.5	343	7	US-10-424-599-51313	Sequence 51313, A
C 432	22	0.6	169659	7	US-10-322-696-70	Sequence 70, Appli	C 505	21	0.5	345	4	US-09-925-065A-732590	Sequence 732590, A
C 433	22	0.6	170245	7	US-10-717-597-322	Sequence 322, App	C 506	21	0.5	353	4	US-09-925-065A-149005	Sequence 149005, A
C 434	22	0.6	176373	5	US-10-095-407-17	Sequence 17, Appli	507	21	0.5	355	3	US-09-867-701-6727	Sequence 6727, Ap
C 435	22	0.6	180557	5	US-10-003-806-6	Sequence 6, Appli	C 508	21	0.5	360	3	US-09-764-877-3230	Sequence 3230, Ap
C 436	22	0.6	249531	9	US-10-003-806-9	Sequence 9, Appli	C 509	21	0.5	360	6	US-10-242-515-3230	Sequence 2336, Ap
C 437	22	0.6	251364	5	US-10-737-318-16	Sequence 16, Appli	510	21	0.5	371	3	US-09-960-352-10509	Sequence 554264, A
C 438	22	0.6	251364	5	US-10-175-523-58	Sequence 58, Appli	511	21	0.5	372	8	US-10-425-115-80509	Sequence 80509, A
C 439	22	0.6	251364	5	US-10-175-523-61	Sequence 61, Appli	512	21	0.5	378	8	US-10-425-115-80509	Sequence 3105, Ap
C 440	22	0.6	251364	5	US-10-175-523-79	Sequence 79, Appli	513	21	0.5	379	3	US-09-960-352-3105	Sequence 938, App
C 441	22	0.6	251364	10	US-11-099-266-58	Sequence 58, Appli	C 514	21	0.5	385	3	US-09-951-945-938	Sequence 938, App
C 442	22	0.6	251364	10	US-11-099-266-61	Sequence 61, Appli	515	21	0.5	385	9	US-10-978-245-938	Sequence 14, Appli
C 443	22	0.6	251364	10	US-11-099-266-79	Sequence 79, Appli	C 516	21	0.5	388	5	US-10-206-901B-14	Sequence 9287, Ap
C 444	22	0.6	260160	8	US-10-723-860-2837	Sequence 2837, Ap	517	21	0.5	390	3	US-09-960-352-9287	Sequence 320, App
C 445	22	0.6	465237	3	US-09-933-267A-1	Sequence 1, Appli	C 518	21	0.5	391	5	US-10-079-623-320	Sequence 314, App
C 446	22	0.6	744802	6	US-10-292-798-1369	Sequence 1369, Ap	C 519	21	0.5	392	5	US-10-079-623-314	Sequence 13149, A
C 447	22	0.6	775062	8	US-10-719-993-6844	Sequence 6844, Ap	520	21	0.5	393	8	US-10-357-930-13149	Sequence 3770, A
C 448	22	0.6	783062	9	US-10-461-862-166	Sequence 166, App	521	21	0.5	396	7	US-10-357-930-37770	Sequence 105806, A
C 449	22	0.6	3673778	6	US-10-312-841-2	Sequence 2, Appli	522	21	0.5	396	7	US-10-424-599-105806	Sequence 177621, A
C 450	21	0.5	21	7	US-10-607-712-43	Sequence 43, Appli	523	21	0.5	398	4	US-09-925-065A-177621	Sequence 177621, A
C 451	21	0.5	21	7	US-10-607-712-43	Sequence 43, Appli	524	21	0.5	403	8	US-10-425-115-61402	Sequence 61402, A
C 452	21	0.5	21	7	US-10-607-712-43	Sequence 43, Appli	525	21	0.5	406	3	US-09-983-965-4402	Sequence 4402, Ap
C 453	21	0.5	21	7	US-10-607-712-43	Sequence 43, Appli	526	21	0.5	410	8	US-10-767-795-977	Sequence 977, App
C 454	21	0.5	39	6	US-10-219-195-37	Sequence 37, Appli	527	21	0.5	410	8	US-10-767-795-977	Sequence 323, App
C 455	21	0.5	141	3	US-09-814-353-1383	Sequence 1383, Ap	C 528	21	0.5	415	5	US-10-079-623-323	Sequence 518970, A
C 456	21	0.5	141	3	US-09-814-353-1383	Sequence 7746, Ap	529	21	0.5	415	4	US-09-925-065A-518970	Sequence 6764, Ap
C 457	21	0.5	170	7	US-10-437-963-56750	Sequence 56750, A	C 530	21	0.5	418	7	US-10-021-323-6764	Sequence 222991, A
C 458	21	0.5	189	5	US-10-079-623-133	Sequence 123, App	531	21	0.5	419	5	US-09-925-065A-222991	Sequence 34810, A
C 459	21	0.5	200	3	US-09-814-353-1387	Sequence 1387, App	C 532	21	0.5	419	5	US-10-027-632-34810	Sequence 34810, A
C 460	21	0.5	200	3	US-09-814-353-1387	Sequence 7750, Ap	C 533	21	0.5	419	6	US-10-027-632-34810	Sequence 136043, A
C 461	21	0.5	201	7	US-10-741-601-6880	Sequence 6880, Ap	534	21	0.5	421	8	US-10-425-115-136043	Sequence 136043, A



C 535	21	0.5	422	8	US-10-357-930-48520	Sequence 48520, A	C 608	21	0.5	511	4	US-09-925-065A-74200	Sequence 74200, A
C 536	21	0.5	424	4	US-09-925-065A-219497	Sequence 219497, A	C 609	21	0.5	511	8	US-10-357-930-34286	Sequence 34286, A
C 537	21	0.5	426	7	US-10-424-599-33658	Sequence 33658, A	C 610	21	0.5	511	8	US-10-357-930-43158	Sequence 43158, A
C 538	21	0.5	427	4	US-09-925-065A-154328	Sequence 154328, A	C 611	21	0.5	512	4	US-09-925-065A-580299	Sequence 580299, A
C 539	21	0.5	427	8	US-10-723-860-7253	Sequence 7253, Ap	C 612	21	0.5	512	4	US-09-925-065A-762931	Sequence 762931, A
C 540	21	0.5	430	7	US-10-424-599-136214	Sequence 136214, A	C 613	21	0.5	512	5	US-10-027-632-267388	Sequence 267388, A
C 541	21	0.5	431	3	US-09-918-995-8709	Sequence 8709, Ap	C 614	21	0.5	512	6	US-10-027-632-267388	Sequence 267388, A
C 542	21	0.5	436	5	US-10-027-632-36674	Sequence 36674, A	C 615	21	0.5	513	4	US-09-925-065A-358459	Sequence 358459, A
C 543	21	0.5	436	5	US-10-027-632-68992	Sequence 68992, A	C 616	21	0.5	513	4	US-09-925-065A-358460	Sequence 358460, A
C 544	21	0.5	436	5	US-10-027-632-69538	Sequence 69538, A	C 617	21	0.5	513	4	US-09-925-065A-358461	Sequence 358461, A
C 545	21	0.5	436	5	US-10-027-632-69538	Sequence 69538, A	C 618	21	0.5	513	4	US-09-925-065A-358462	Sequence 358462, A
C 546	21	0.5	436	6	US-10-027-632-36674	Sequence 36674, A	C 619	21	0.5	513	4	US-09-925-065A-763143	Sequence 763143, A
C 547	21	0.5	436	6	US-10-027-632-68992	Sequence 68992, A	C 620	21	0.5	514	4	US-09-925-065A-484457	Sequence 484457, A
C 548	21	0.5	436	6	US-10-027-632-69538	Sequence 69538, A	C 621	21	0.5	516	5	US-10-027-632-237042	Sequence 237042, A
C 549	21	0.5	436	6	US-10-027-632-311766	Sequence 311766, A	C 622	21	0.5	516	6	US-10-027-632-237042	Sequence 237042, A
C 550	21	0.5	437	6	US-10-264-049-1763	Sequence 1763, Ap	C 623	21	0.5	521	4	US-09-925-065A-422732	Sequence 422732, A
C 551	21	0.5	442	7	US-10-021-323-16032	Sequence 16032, A	C 624	21	0.5	521	4	US-09-814-353-14057	Sequence 14057, A
C 552	21	0.5	442	8	US-10-357-930-10102	Sequence 10102, A	C 625	21	0.5	524	3	US-09-925-065A-118032	Sequence 118032, A
C 553	21	0.5	444	3	US-09-814-353-14135	Sequence 14135, A	C 626	21	0.5	525	7	US-10-021-323-3074	Sequence 3074, Ap
C 554	21	0.5	444	5	US-10-027-632-63810	Sequence 63810, A	C 627	21	0.5	525	8	US-10-357-930-56012	Sequence 56012, A
C 555	21	0.5	444	5	US-10-027-632-63811	Sequence 63811, A	C 628	21	0.5	526	4	US-09-925-065A-320969	Sequence 320969, A
C 556	21	0.5	444	6	US-10-027-632-63811	Sequence 63811, A	C 629	21	0.5	526	4	US-09-925-065A-320970	Sequence 320970, A
C 557	21	0.5	444	6	US-10-027-632-63811	Sequence 63811, A	C 630	21	0.5	529	7	US-10-437-963-44744	Sequence 44744, A
C 558	21	0.5	444	8	US-10-357-930-49422	Sequence 49422, A	C 631	21	0.5	530	4	US-09-925-065A-318951	Sequence 318951, A
C 559	21	0.5	445	4	US-09-925-065A-349522	Sequence 349522, A	C 632	21	0.5	530	4	US-09-925-065A-318953	Sequence 318953, A
C 560	21	0.5	446	5	US-10-027-632-281108	Sequence 281108, A	C 633	21	0.5	530	4	US-09-925-065A-318954	Sequence 318954, A
C 561	21	0.5	446	5	US-10-027-632-281109	Sequence 281109, A	C 634	21	0.5	530	4	US-09-925-065A-557580	Sequence 557580, A
C 562	21	0.5	446	5	US-10-027-632-281110	Sequence 281110, A	C 635	21	0.5	532	4	US-09-925-065A-610344	Sequence 610344, A
C 563	21	0.5	446	6	US-10-027-632-281108	Sequence 281108, A	C 636	21	0.5	532	4	US-09-925-065A-610345	Sequence 610345, A
C 564	21	0.5	446	6	US-10-027-632-281109	Sequence 281109, A	C 637	21	0.5	532	4	US-09-925-065A-610346	Sequence 610346, A
C 565	21	0.5	446	6	US-10-027-632-281110	Sequence 281110, A	C 638	21	0.5	532	4	US-09-925-065A-610347	Sequence 610347, A
C 566	21	0.5	455	3	US-09-833-790-69	Sequence 69, Appl	C 639	21	0.5	533	4	US-09-925-065A-526099	Sequence 526099, A
C 567	21	0.5	456	4	US-09-925-065A-513522	Sequence 513522, A	C 640	21	0.5	533	5	US-10-027-632-197539	Sequence 197539, A
C 568	21	0.5	456	4	US-09-925-065A-561045	Sequence 561045, A	C 641	21	0.5	533	6	US-10-027-632-197539	Sequence 197539, A
C 569	21	0.5	456	4	US-09-925-065A-561047	Sequence 561047, A	C 642	21	0.5	534	4	US-09-925-065A-10121	Sequence 10121, A
C 570	21	0.5	456	4	US-09-925-065A-561048	Sequence 561048, A	C 643	21	0.5	534	4	US-09-925-065A-333898	Sequence 333898, A
C 571	21	0.5	457	3	US-09-864-761-1586	Sequence 1586, Ap	C 644	21	0.5	537	4	US-09-925-065A-230349	Sequence 230349, A
C 572	21	0.5	463	7	US-10-437-963-7814	Sequence 7814, Ap	C 645	21	0.5	537	4	US-09-925-065A-489872	Sequence 489872, A
C 573	21	0.5	466	7	US-10-424-599-142436	Sequence 142436, Ap	C 646	21	0.5	537	4	US-09-925-065A-489873	Sequence 489873, A
C 574	21	0.5	467	8	US-10-357-930-933	Sequence 933, App	C 647	21	0.5	537	4	US-09-925-065A-489874	Sequence 489874, A
C 575	21	0.5	470	7	US-10-424-599-92797	Sequence 92797, A	C 648	21	0.5	538	4	US-09-925-065A-340190	Sequence 340190, A
C 576	21	0.5	471	4	US-09-925-065A-496565	Sequence 496565, A	C 649	21	0.5	542	5	US-10-027-632-73650	Sequence 73650, A
C 577	21	0.5	471	4	US-09-925-065A-496566	Sequence 496566, A	C 650	21	0.5	542	5	US-10-027-632-73651	Sequence 73651, A
C 578	21	0.5	472	3	US-09-814-353-14131	Sequence 14131, A	C 651	21	0.5	542	5	US-10-027-632-74541	Sequence 74541, A
C 579	21	0.5	472	5	US-10-066-543-1608	Sequence 1608, Ap	C 652	21	0.5	542	5	US-10-027-632-74542	Sequence 74542, A
C 580	21	0.5	476	5	US-10-027-632-282965	Sequence 282965, A	C 653	21	0.5	542	5	US-10-027-632-299183	Sequence 299183, A
C 581	21	0.5	476	6	US-10-027-632-282965	Sequence 282965, A	C 654	21	0.5	542	5	US-10-027-632-299184	Sequence 299184, A
C 582	21	0.5	478	3	US-09-864-761-11552	Sequence 11552, A	C 655	21	0.5	542	6	US-10-027-632-73650	Sequence 73650, A
C 583	21	0.5	482	5	US-10-198-846-1141	Sequence 1141, Ap	C 656	21	0.5	542	6	US-10-027-632-73651	Sequence 73651, A
C 584	21	0.5	485	6	US-10-058-053A-220	Sequence 220, App	C 657	21	0.5	542	6	US-10-027-632-74541	Sequence 74541, A
C 585	21	0.5	485	8	US-10-838-226-220	Sequence 220, App	C 658	21	0.5	542	6	US-10-027-632-74542	Sequence 74542, A
C 586	21	0.5	486	7	US-10-424-599-50782	Sequence 50782, A	C 659	21	0.5	542	6	US-10-027-632-299183	Sequence 299183, A
C 587	21	0.5	486	8	US-10-357-930-31301	Sequence 31301, A	C 660	21	0.5	546	4	US-09-925-065A-71178	Sequence 71178, A
C 588	21	0.5	486	8	US-10-357-930-40271	Sequence 40271, A	C 661	21	0.5	546	4	US-09-925-065A-71179	Sequence 71179, A
C 589	21	0.5	489	4	US-09-925-065A-289154	Sequence 289154, A	C 662	21	0.5	547	4	US-09-925-065A-589347	Sequence 589347, A
C 590	21	0.5	489	5	US-10-027-632-37180	Sequence 37180, A	C 663	21	0.5	550	4	US-09-925-065A-319926	Sequence 319926, A
C 591	21	0.5	489	6	US-10-027-632-37180	Sequence 37180, A	C 664	21	0.5	550	4	US-09-925-065A-319927	Sequence 319927, A
C 592	21	0.5	493	8	US-10-357-930-40842	Sequence 40842, A	C 665	21	0.5	550	4	US-09-925-065A-591297	Sequence 591297, A
C 593	21	0.5	496	4	US-09-925-065A-753404	Sequence 753404, A	C 666	21	0.5	556	4	US-09-925-065A-193387	Sequence 193387, A
C 594	21	0.5	496	4	US-09-925-065A-753405	Sequence 753405, A	C 667	21	0.5	557	4	US-09-925-065A-293387	Sequence 293387, A
C 595	21	0.5	497	4	US-09-925-065A-260244	Sequence 260244, A	C 668	21	0.5	558	4	US-09-925-065A-454776	Sequence 454776, A
C 596	21	0.5	498	7	US-10-424-599-56948	Sequence 56948, A	C 669	21	0.5	558	4	US-09-925-065A-454777	Sequence 454777, A
C 597	21	0.5	500	4	US-09-925-065A-252253	Sequence 252253, A	C 670	21	0.5	558	4	US-10-437-963-63844	Sequence 63844, A
C 598	21	0.5	500	5	US-10-060-830-40	Sequence 40, Appl	C 671	21	0.5	559	7	US-09-925-065A-17376	Sequence 17376, A
C 599	21	0.5	501	8	US-10-357-930-20613	Sequence 20613, A	C 672	21	0.5	560	4	US-10-027-632-283307	Sequence 283307, A
C 600	21	0.5	501	8	US-10-357-930-26452	Sequence 26452, A	C 673	21	0.5	561	5	US-10-027-632-283307	Sequence 283307, A
C 601	21	0.5	503	4	US-09-925-065A-387021	Sequence 387021, A	C 674	21	0.5	562	4	US-09-925-065A-158888	Sequence 158888, A
C 602	21	0.5	503	4	US-09-925-065A-387022	Sequence 387022, A	C 675	21	0.5	562	4	US-10-027-632-288589	Sequence 288589, A
C 603	21	0.5	505	4	US-09-925-065A-751623	Sequence 751623, A	C 676	21	0.5	565	5	US-10-027-632-288589	Sequence 288589, A
C 604	21	0.5	510	4	US-09-925-065A-405485	Sequence 405485, A	C 677	21	0.5	565	5	US-10-027-632-308642	Sequence 308642, A
C 605	21	0.5	510	4	US-09-925-065A-405486	Sequence 405486, A	C 678	21	0.5	566	5	US-10-027-632-308642	Sequence 308642, A
C 606	21	0.5	511	4	US-09-925-065A-74198	Sequence 74198, A	C 679	21	0.5	566	6	US-10-027-632-1053261	Sequence 1053261, A
C 607	21	0.5	511	4	US-09-925-065A-74199	Sequence 74199, A	C 680	21	0.5	571	8	US-10-425-115-153261	Sequence 153261, A

681	21	0.5	572	5	US-10-027-632-266149	Sequence 266149,	754	21	0.5	612	4	US-09-925-065A-343174	Sequence 343174,
682	21	0.5	572	6	US-10-027-632-266150	Sequence 266150,	755	21	0.5	612	4	US-09-925-065A-343175	Sequence 343175,
683	21	0.5	572	5	US-10-027-632-266149	Sequence 266149,	756	21	0.5	612	4	US-09-925-065A-343176	Sequence 343176,
684	21	0.5	572	6	US-10-027-632-266150	Sequence 266150,	c 757	21	0.5	613	4	US-09-925-065A-928668	Sequence 928668,
685	21	0.5	575	4	US-09-925-065A-938868	Sequence 938868,	c 758	21	0.5	613	4	US-09-925-065A-770572	Sequence 770572,
686	21	0.5	575	4	US-09-925-065A-938869	Sequence 938869,	c 759	21	0.5	613	4	US-09-925-065A-47436	Sequence 47436,
687	21	0.5	577	4	US-09-925-065A-258961	Sequence 258961,	c 760	21	0.5	619	4	US-09-925-065A-56506	Sequence 56506,
688	21	0.5	577	4	US-09-925-065A-258962	Sequence 258962,	c 761	21	0.5	620	4	US-09-925-065A-866942	Sequence 866942,
689	21	0.5	577	9	US-10-956-157-3052	Sequence 3052, Ap	c 762	21	0.5	620	4	US-09-925-065A-866943	Sequence 866943,
690	21	0.5	577	9	US-10-956-157-3052	Sequence 3052, Ap	c 763	21	0.5	620	4	US-09-925-065A-899725	Sequence 899725,
691	21	0.5	579	4	US-09-925-065A-954549	Sequence 954549,	c 764	21	0.5	622	4	US-09-925-065A-501311	Sequence 501311,
692	21	0.5	580	4	US-09-925-065A-296463	Sequence 296463,	c 765	21	0.5	622	4	US-09-925-065A-501312	Sequence 501312,
693	21	0.5	582	4	US-09-925-065A-899637	Sequence 899637,	c 766	21	0.5	622	7	US-10-767-701-19690	Sequence 19690,
694	21	0.5	582	4	US-09-925-065A-899638	Sequence 899638,	c 767	21	0.5	623	4	US-09-925-065A-264948	Sequence 264948,
695	21	0.5	583	4	US-09-925-065A-40758	Sequence 40758, A	c 768	21	0.5	624	5	US-10-027-632-294166	Sequence 294166,
696	21	0.5	583	4	US-09-925-065A-558657	Sequence 558657,	c 769	21	0.5	624	6	US-10-027-632-294167	Sequence 294167,
697	21	0.5	585	3	US-09-814-353-1308	Sequence 1308, Ap	c 770	21	0.5	624	6	US-10-027-632-294166	Sequence 294166,
698	21	0.5	585	3	US-09-814-353-1308	Sequence 1308, Ap	c 771	21	0.5	624	6	US-10-027-632-294166	Sequence 294166,
699	21	0.5	585	4	US-09-925-065A-172098	Sequence 172098,	c 772	21	0.5	625	4	US-09-925-065A-15616	Sequence 15616,
700	21	0.5	585	4	US-09-925-065A-172099	Sequence 172099,	c 773	21	0.5	625	4	US-09-925-065A-158695	Sequence 158695,
701	21	0.5	585	4	US-09-925-065A-172100	Sequence 172100,	c 774	21	0.5	625	4	US-09-925-065A-158696	Sequence 158696,
702	21	0.5	585	4	US-09-925-065A-172100	Sequence 172100,	c 775	21	0.5	626	4	US-09-925-065A-938813	Sequence 938813,
703	21	0.5	585	4	US-09-925-065A-294385	Sequence 294385,	c 776	21	0.5	626	4	US-09-925-065A-938814	Sequence 938814,
704	21	0.5	585	4	US-09-925-065A-294386	Sequence 294386,	c 777	21	0.5	627	4	US-09-925-065A-878743	Sequence 878743,
705	21	0.5	585	4	US-09-925-065A-774157	Sequence 774157,	c 778	21	0.5	627	4	US-09-925-065A-878744	Sequence 878744,
706	21	0.5	587	4	US-09-925-065A-670881	Sequence 670881,	c 779	21	0.5	627	4	US-09-925-065A-878745	Sequence 878745,
707	21	0.5	587	4	US-09-925-065A-670882	Sequence 670882,	c 780	21	0.5	628	4	US-09-925-065A-559335	Sequence 559335,
708	21	0.5	587	4	US-09-925-065A-670883	Sequence 670883,	c 781	21	0.5	629	5	US-10-027-632-54606	Sequence 54606,
709	21	0.5	587	5	US-10-027-632-214640	Sequence 214640,	c 782	21	0.5	629	5	US-10-027-632-54607	Sequence 54607,
710	21	0.5	587	6	US-10-027-632-214640	Sequence 214640,	c 783	21	0.5	629	6	US-10-027-632-54607	Sequence 54607,
711	21	0.5	587	8	US-10-357-930-28715	Sequence 28715, A	c 784	21	0.5	629	6	US-10-027-632-54607	Sequence 54607,
712	21	0.5	587	8	US-10-357-930-28715	Sequence 28715, A	c 785	21	0.5	631	7	US-10-027-632-54607	Sequence 54607,
713	21	0.5	589	4	US-09-925-065A-588995	Sequence 588995,	c 786	21	0.5	632	5	US-10-027-632-77235	Sequence 77235,
714	21	0.5	589	4	US-09-925-065A-588996	Sequence 588996,	c 787	21	0.5	632	5	US-10-027-632-77235	Sequence 77235,
715	21	0.5	589	4	US-09-925-065A-588997	Sequence 588997,	c 788	21	0.5	632	6	US-10-027-632-77235	Sequence 77235,
716	21	0.5	589	4	US-09-925-065A-588998	Sequence 588998,	c 789	21	0.5	633	6	US-10-027-632-77235	Sequence 77235,
717	21	0.5	592	4	US-09-925-065A-895820	Sequence 895820,	c 790	21	0.5	633	4	US-09-925-065A-875899	Sequence 875899,
718	21	0.5	592	5	US-10-027-632-264193	Sequence 264193,	c 791	21	0.5	636	4	US-09-925-065A-731134	Sequence 731134,
719	21	0.5	592	5	US-10-027-632-264193	Sequence 264193,	c 792	21	0.5	636	4	US-09-925-065A-731135	Sequence 731135,
720	21	0.5	592	6	US-10-027-632-264193	Sequence 264193,	c 793	21	0.5	636	5	US-10-027-632-423	Sequence 423, App
721	21	0.5	592	6	US-10-027-632-264193	Sequence 264193,	c 794	21	0.5	636	5	US-10-027-632-261747	Sequence 261747,
722	21	0.5	592	6	US-10-027-632-264193	Sequence 264193,	c 795	21	0.5	636	5	US-10-027-632-261748	Sequence 261748,
723	21	0.5	593	4	US-09-925-065A-472185	Sequence 472185,	c 796	21	0.5	636	5	US-10-198-846-8095	Sequence 8095, App
724	21	0.5	593	4	US-09-925-065A-866355	Sequence 866355,	c 797	21	0.5	636	6	US-10-027-632-423	Sequence 423, App
725	21	0.5	595	4	US-09-925-065A-911361	Sequence 911361,	c 798	21	0.5	636	6	US-10-027-632-261747	Sequence 261747,
726	21	0.5	597	4	US-09-925-065A-762330	Sequence 762330,	c 799	21	0.5	636	6	US-10-027-632-261748	Sequence 261748,
727	21	0.5	597	8	US-10-468-488-32	Sequence 32, App1	c 800	21	0.5	637	4	US-09-925-065A-932840	Sequence 932840,
728	21	0.5	598	4	US-09-925-065A-816088	Sequence 816088,	c 801	21	0.5	637	4	US-09-925-065A-548422	Sequence 548422,
729	21	0.5	599	4	US-09-925-065A-399183	Sequence 399183,	c 802	21	0.5	638	4	US-09-925-065A-794334	Sequence 794334,
730	21	0.5	599	4	US-09-925-065A-399184	Sequence 399184,	c 803	21	0.5	639	4	US-09-925-065A-822586	Sequence 822586,
731	21	0.5	599	9	US-10-973-079-23469	Sequence 23469, A	c 804	21	0.5	640	4	US-09-925-065A-688042	Sequence 688042,
732	21	0.5	600	4	US-09-925-065A-690628	Sequence 690628,	c 805	21	0.5	640	4	US-09-925-065A-688043	Sequence 688043,
733	21	0.5	600	9	US-10-956-157-6252	Sequence 6252, Ap	c 806	21	0.5	640	4	US-09-925-065A-862829	Sequence 862829,
734	21	0.5	600	9	US-10-956-157-7104	Sequence 7104, Ap	c 807	21	0.5	640	4	US-09-925-065A-862830	Sequence 862830,
735	21	0.5	600	9	US-10-972-079-73100	Sequence 73100, A	c 808	21	0.5	640	4	US-09-925-065A-897747	Sequence 897747,
736	21	0.5	600	9	US-10-972-079-73100	Sequence 73100, A	c 809	21	0.5	645	3	US-09-923-246-108	Sequence 108, App
737	21	0.5	600	10	US-11-060-756-2344	Sequence 2344, Ap	c 810	21	0.5	645	7	US-10-295-723-108	Sequence 108, App
738	21	0.5	600	10	US-11-060-756-2345	Sequence 2345, Ap	c 811	21	0.5	645	7	US-10-559-684-108	Sequence 108, App
739	21	0.5	600	10	US-11-060-756-2874	Sequence 2874, Ap	c 812	21	0.5	645	8	US-10-787-442-108	Sequence 108, App
740	21	0.5	600	10	US-11-060-756-6616	Sequence 6616, Ap	c 813	21	0.5	649	5	US-10-027-632-212510	Sequence 212510,
741	21	0.5	600	10	US-11-060-756-6617	Sequence 6617, Ap	c 814	21	0.5	649	5	US-10-027-632-212511	Sequence 212511,
742	21	0.5	600	10	US-11-060-756-7146	Sequence 7146, Ap	c 815	21	0.5	649	6	US-10-027-632-212510	Sequence 212510,
743	21	0.5	603	4	US-09-925-065A-319246	Sequence 319246,	c 816	21	0.5	649	6	US-10-027-632-212511	Sequence 212511,
744	21	0.5	610	4	US-09-925-065A-580587	Sequence 580587,	c 817	21	0.5	650	4	US-09-925-065A-701394	Sequence 701394,
745	21	0.5	611	4	US-09-925-065A-885141	Sequence 885141,	c 818	21	0.5	653	4	US-09-925-065A-748516	Sequence 748516,
746	21	0.5	611	4	US-09-925-065A-910826	Sequence 910826,	c 819	21	0.5	653	4	US-09-925-065A-748517	Sequence 748517,
747	21	0.5	611	5	US-10-027-632-195597	Sequence 195597,	c 820	21	0.5	653	4	US-09-925-065A-821566	Sequence 821566,
748	21	0.5	611	5	US-10-027-632-195598	Sequence 195598,	c 821	21	0.5	657	4	US-09-925-065A-41772	Sequence 41772, A
749	21	0.5	611	6	US-10-027-632-195597	Sequence 195597,	c 822	21	0.5	657	4	US-09-925-065A-41773	Sequence 41773, A
750	21	0.5	611	6	US-10-027-632-195598	Sequence 195598,	c 823	21	0.5	657	8	US-10-425-115-179521	Sequence 179521,
751	21	0.5	612	4	US-09-925-065A-161287	Sequence 161287,	c 824	21	0.5	668	4	US-09-925-065A-691735	Sequence 691735,
752	21	0.5	612	4	US-09-925-065A-161288	Sequence 161288,	c 825	21	0.5	672	4	US-09-925-065A-682472	Sequence 682472,
753	21	0.5	612	4	US-09-925-065A-247900	Sequence 247900,	c 826	21	0.5	673	4	US-09-925-065A-680053	Sequence 680053,

C 827	21	0.5	674	3	US-09-825-301-21	Sequence 21, Appl	900	21	0.5	913	6	US-10-227-577-1662	Sequence 1662, Ap
C 828	21	0.5	674	3	US-09-604-287A-465	Sequence 465, App	901	21	0.5	914	3	US-09-764-869-1661	Sequence 1661, Ap
C 829	21	0.5	674	3	US-09-834-759-465	Sequence 465, App	C 902	21	0.5	914	5	US-10-027-632-122423	Sequence 122423, Ap
C 830	21	0.5	674	3	US-09-551-621-465	Sequence 465, App	C 903	21	0.5	914	5	US-10-027-632-122424	Sequence 122424, Ap
C 831	21	0.5	674	5	US-10-007-805-465	Sequence 465, App	C 904	21	0.5	914	5	US-10-091-504-1661	Sequence 1661, Ap
C 832	21	0.5	674	5	US-10-076-802-465	Sequence 465, App	C 905	21	0.5	914	6	US-10-027-632-122423	Sequence 122423, Ap
C 833	21	0.5	674	6	US-10-124-805-465	Sequence 465, App	C 906	21	0.5	914	6	US-10-027-632-122424	Sequence 122424, Ap
C 834	21	0.5	674	6	US-10-033-527-21	Sequence 21, Appl	C 907	21	0.5	914	6	US-10-227-577-1661	Sequence 1661, Ap
C 835	21	0.5	678	8	US-10-723-860-2914	Sequence 2914, Ap	C 908	21	0.5	918	5	US-10-027-632-164199	Sequence 164199, Ap
C 836	21	0.5	678	8	US-10-756-149-2707	Sequence 2707, Ap	C 909	21	0.5	918	6	US-10-027-632-164199	Sequence 164199, Ap
C 837	21	0.5	679	8	US-10-653-047-7211	Sequence 7211, Ap	C 910	21	0.5	918	7	US-10-264-237-682	Sequence 682, App
C 838	21	0.5	685	5	US-10-027-632-89169	Sequence 89169, A	C 911	21	0.5	918	7	US-10-437-963-90521	Sequence 90521, A
C 839	21	0.5	685	5	US-10-027-632-89170	Sequence 89170, A	C 912	21	0.5	923	9	US-10-773-236-37	Sequence 37, Appl
C 840	21	0.5	685	6	US-10-027-632-89169	Sequence 89169, A	C 913	21	0.5	931	6	US-10-074-024-636	Sequence 636, App
C 841	21	0.5	685	6	US-10-027-632-89170	Sequence 89170, A	C 914	21	0.5	931	6	US-10-074-024-637	Sequence 637, App
C 842	21	0.5	690	5	US-10-027-632-135875	Sequence 135875, A	C 915	21	0.5	961	5	US-10-027-632-33420	Sequence 33420, A
C 843	21	0.5	690	6	US-10-027-632-135875	Sequence 135875, A	C 916	21	0.5	961	6	US-10-027-632-33420	Sequence 33420, A
C 844	21	0.5	700	4	US-09-925-065A-540081	Sequence 540081, A	C 917	21	0.5	994	5	US-10-027-632-31531	Sequence 31531, A
C 845	21	0.5	710	4	US-09-925-065A-540081	Sequence 540081, A	C 918	21	0.5	994	6	US-10-027-632-31531	Sequence 31531, A
C 846	21	0.5	725	5	US-10-027-632-149502	Sequence 149502, A	C 919	21	0.5	998	8	US-10-739-930-1855	Sequence 1855, Ap
C 847	21	0.5	725	5	US-10-027-632-149502	Sequence 149502, A	C 920	21	0.5	999	6	US-10-355-430-23	Sequence 23, Appl
C 848	21	0.5	725	5	US-10-027-632-149504	Sequence 149504, A	C 921	21	0.5	1002	4	US-09-925-065A-31170	Sequence 31170, A
C 849	21	0.5	725	5	US-10-027-632-149504	Sequence 149504, A	C 922	21	0.5	1027	4	US-09-925-065A-3659	Sequence 3659, Ap
C 850	21	0.5	725	6	US-10-027-632-149502	Sequence 149502, A	C 923	21	0.5	1027	4	US-09-925-065A-3660	Sequence 3660, Ap
C 851	21	0.5	725	6	US-10-027-632-149502	Sequence 149502, A	C 924	21	0.5	1027	4	US-09-925-065A-3661	Sequence 3661, Ap
C 852	21	0.5	725	6	US-10-027-632-149504	Sequence 149504, A	C 925	21	0.5	1027	4	US-09-925-065A-3662	Sequence 3662, Ap
C 853	21	0.5	725	6	US-10-027-632-149504	Sequence 149504, A	C 926	21	0.5	1027	4	US-09-925-065A-3662	Sequence 3662, Ap
C 854	21	0.5	726	4	US-09-925-065A-63027	Sequence 63027, A	C 927	21	0.5	1051	8	US-10-425-115-123330	Sequence 123330, Ap
C 855	21	0.5	727	5	US-10-027-632-28373	Sequence 28373, A	C 928	21	0.5	1070	4	US-09-925-065A-686010	Sequence 686010, A
C 856	21	0.5	727	6	US-10-027-632-28373	Sequence 28373, A	C 929	21	0.5	1070	4	US-09-925-065A-686011	Sequence 686011, A
C 857	21	0.5	732	4	US-10-027-632-28373	Sequence 28373, A	C 930	21	0.5	1087	4	US-09-925-065A-709482	Sequence 709482, Ap
C 858	21	0.5	735	9	US-10-489-740-70	Sequence 70, Appl	C 931	21	0.5	1087	4	US-09-925-065A-709483	Sequence 709483, Ap
C 859	21	0.5	744	5	US-10-027-632-110802	Sequence 110802, A	C 932	21	0.5	1087	4	US-09-925-065A-709483	Sequence 709483, Ap
C 860	21	0.5	744	6	US-10-027-632-110802	Sequence 110802, A	C 933	21	0.5	1092	4	US-09-925-065A-689194	Sequence 689194, Ap
C 861	21	0.5	762	4	US-09-925-065A-59562	Sequence 59562, A	C 934	21	0.5	1130	4	US-09-925-065A-719120	Sequence 719120, Ap
C 862	21	0.5	762	4	US-09-925-065A-59562	Sequence 59562, A	C 935	21	0.5	1130	4	US-09-925-065A-719121	Sequence 719121, Ap
C 863	21	0.5	764	5	US-09-925-065A-59563	Sequence 59563, A	C 936	21	0.5	1131	4	US-09-925-065A-718023	Sequence 718023, Ap
C 864	21	0.5	764	6	US-10-027-632-123808	Sequence 123808, A	C 937	21	0.5	1139	4	US-09-925-065A-718024	Sequence 718024, Ap
C 865	21	0.5	772	8	US-10-027-632-123808	Sequence 123808, A	C 938	21	0.5	1139	4	US-09-925-065A-28729	Sequence 28729, A
C 866	21	0.5	772	8	US-10-653-047-7581	Sequence 7581, Ap	C 939	21	0.5	1139	4	US-09-925-065A-28730	Sequence 28730, A
C 867	21	0.5	774	5	US-10-027-632-143728	Sequence 143728, A	C 940	21	0.5	1139	4	US-09-925-065A-28731	Sequence 28731, A
C 868	21	0.5	777	8	US-10-027-632-143728	Sequence 143728, A	C 941	21	0.5	1139	4	US-09-925-065A-28732	Sequence 28732, A
C 869	21	0.5	781	7	US-10-425-115-123862	Sequence 123862, A	C 942	21	0.5	1167	4	US-09-925-065A-670666	Sequence 670666, Ap
C 870	21	0.5	782	7	US-10-767-701-12537	Sequence 12537, A	C 943	21	0.5	1214	8	US-09-925-065A-704144	Sequence 704144, Ap
C 871	21	0.5	782	7	US-10-437-963-33683	Sequence 33683, A	C 944	21	0.5	1243	5	US-10-424-599-50500	Sequence 50500, A
C 872	21	0.5	797	5	US-10-027-632-144701	Sequence 144701, A	C 945	21	0.5	1252	7	US-10-424-599-50500	Sequence 50500, A
C 873	21	0.5	797	6	US-10-027-632-144701	Sequence 144701, A	C 946	21	0.5	1284	4	US-09-925-065A-681170	Sequence 681170, Ap
C 874	21	0.5	798	5	US-10-027-632-30877	Sequence 30877, A	C 947	21	0.5	1284	4	US-09-925-065A-681171	Sequence 681171, Ap
C 875	21	0.5	798	6	US-10-027-632-30877	Sequence 30877, A	C 948	21	0.5	1284	4	US-09-925-065A-681172	Sequence 681172, Ap
C 876	21	0.5	801	4	US-09-925-065A-678775	Sequence 678775, A	C 949	21	0.5	1284	4	US-10-240-425-1567	Sequence 1567, Ap
C 877	21	0.5	801	5	US-10-027-632-131524	Sequence 131524, A	C 950	21	0.5	1292	7	US-10-363-616-30	Sequence 30, Appl
C 878	21	0.5	801	5	US-10-027-632-131524	Sequence 131524, A	C 951	21	0.5	1312	6	US-10-133-013-131	Sequence 131, Appl
C 879	21	0.5	801	5	US-10-027-632-131526	Sequence 131526, A	C 952	21	0.5	1318	3	US-09-921-406C-7	Sequence 7, Appl
C 880	21	0.5	801	5	US-10-027-632-131526	Sequence 131526, A	C 953	21	0.5	1318	6	US-10-172-118-938	Sequence 938, App
C 881	21	0.5	801	6	US-10-027-632-131524	Sequence 131524, A	C 954	21	0.5	1318	6	US-10-159-563-151	Sequence 938, App
C 882	21	0.5	801	6	US-10-027-632-131524	Sequence 131524, A	C 955	21	0.5	1318	7	US-10-342-887-938	Sequence 938, App
C 883	21	0.5	801	6	US-10-027-632-131526	Sequence 131526, A	C 956	21	0.5	1318	7	US-10-755-889-391	Sequence 391, App
C 884	21	0.5	801	6	US-10-027-632-131526	Sequence 131526, A	C 957	21	0.5	1318	9	US-10-848-755A-117	Sequence 117, App
C 885	21	0.5	807	9	US-10-479-081-327	Sequence 327, App	C 958	21	0.5	1331	4	US-09-925-065A-47638	Sequence 47638, A
C 886	21	0.5	824	5	US-10-027-632-156748	Sequence 156748, A	C 959	21	0.5	1342	3	US-09-925-297-50	Sequence 50, Appl
C 887	21	0.5	824	6	US-10-027-632-156748	Sequence 156748, A	C 960	21	0.5	1382	4	US-09-925-065A-69581	Sequence 69581, A
C 888	21	0.5	836	7	US-10-437-963-33817	Sequence 33817, A	C 961	21	0.5	1382	4	US-09-925-065A-69582	Sequence 69582, A
C 889	21	0.5	838	5	US-10-027-632-148806	Sequence 148806, A	C 962	21	0.5	1382	4	US-09-925-065A-69583	Sequence 69583, A
C 890	21	0.5	838	6	US-10-027-632-148806	Sequence 148806, A	C 963	21	0.5	1382	4	US-09-925-065A-69584	Sequence 69584, A
C 891	21	0.5	873	7	US-10-424-599-4989	Sequence 4989, Ap	C 964	21	0.5	1382	4	US-09-925-065A-550969	Sequence 550969, Ap
C 892	21	0.5	883	7	US-10-767-701-1273	Sequence 1273, Ap	C 965	21	0.5	1387	5	US-10-027-632-250056	Sequence 250056, Ap
C 893	21	0.5	896	5	US-10-106-698-188	Sequence 188, App	C 966	21	0.5	1387	6	US-10-027-632-250056	Sequence 250056, Ap
C 894	21	0.5	907	4	US-09-925-065A-68283	Sequence 68283, A	C 967	21	0.5	1512	4	US-09-925-065A-704680	Sequence 704680, Ap
C 895	21	0.5	907	4	US-09-925-065A-68284	Sequence 68284, A	C 968	21	0.5	1512	4	US-09-925-065A-704681	Sequence 704681, Ap
C 896	21	0.5	907	4	US-09-925-065A-68285	Sequence 68285, A	C 969	21	0.5	1521	6	US-10-172-118-1450	Sequence 1450, Ap
C 897	21	0.5	908	8	US-10-357-930-7571	Sequence 7571, Ap	C 970	21	0.5	1521	7	US-10-342-887-1450	Sequence 1450, Ap
C 898	21	0.5	913	3	US-09-764-869-1662	Sequence 1662, Ap	C 971	21	0.5	1521	9	US-10-956-157-1869	Sequence 1869, Ap
C 899	21	0.5	913	5	US-10-091-504-1662	Sequence 1662, Ap	C 972	21	0.5	1559	7	US-10-351-334-42	Sequence 42, Appl

c 973 21 0.5 1569 9 US-10-450-763-19923 Sequence 19923, A  
c 974 21 0.5 1580 6 US-10-264-049-144 Sequence 144, App  
c 975 21 0.5 1583 5 US-10-000-256A-32 Sequence 32, Appl  
c 976 21 0.5 1639 3 US-09-731-872-6 Sequence 6, Appli  
c 977 21 0.5 1639 3 US-09-876-997-6 Sequence 6, Appli  
c 978 21 0.5 1639 9 US-10-643-836-6 Sequence 2486, Ap  
c 979 21 0.5 1645 8 US-10-739-930-2486 Sequence 11419, A  
c 980 21 0.5 1648 8 US-10-425-115-11419 Sequence 73000, A  
c 981 21 0.5 1673 4 US-09-925-065A-73000 Sequence 47, Appl  
c 982 21 0.5 1674 7 US-10-439-703-47 Sequence 14, Appl  
c 983 21 0.5 1681 3 US-09-997-003-14 Sequence 699512, A  
c 984 21 0.5 1694 4 US-09-925-065A-699512 Sequence 699513, A  
c 985 21 0.5 1694 4 US-09-925-065A-699513 Sequence 714280, A  
c 986 21 0.5 1777 5 US-09-925-065A-714280 Sequence 266309, A  
c 987 21 0.5 1777 5 US-10-027-632-266309 Sequence 266309, A  
c 988 21 0.5 1777 6 US-10-027-632-266309 Sequence 54, Appl  
c 989 21 0.5 1778 3 US-09-791-932-54 Sequence 685610, A  
c 990 21 0.5 1798 4 US-09-925-065A-685610 Sequence 99461, A  
c 991 21 0.5 1801 8 US-10-425-115-99461 Sequence 99461, A  
c 992 21 0.5 1803 8 US-10-723-860-9969 Sequence 954, App  
c 993 21 0.5 1829 5 US-10-106-698-954 Sequence 4110, Ap  
c 994 21 0.5 1853 8 US-10-739-930-4110 Sequence 22230, A  
c 995 21 0.5 1868 8 US-10-357-930-22230 Sequence 28088, A  
c 996 21 0.5 1868 8 US-10-357-930-28088 Sequence 14, Appl  
c 997 21 0.5 1918 5 US-10-036-542-14 Sequence 38, Appl  
c 998 21 0.5 1927 5 US-10-036-542-38 Sequence 132923, A  
c 999 21 0.5 1990 8 US-10-425-115-132923 Sequence 24741, A  
c1000 21 0.5 2023 8 US-10-357-930-24741

ALIGNMENTS

RESULT 1  
US-10-607-712-1  
; Sequence 1, Application US/10607712  
; Publication No. US20040073018A1  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER  
; FILE REFERENCE: 50004/002005  
; CURRENT APPLICATION NUMBER: US/10/607,712  
; CURRENT FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 08/980,326  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/031,964  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: 60/050,310  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3919  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Other  
; LOCATION: (1)...(3919)  
; OTHER INFORMATION: Entire cloned cDNA encoding wild type methionine  
; OTHER INFORMATION: synthase.  
US-10-607-712-1  
Query Match 100.0%; Score 3919; DB 7; Length 3919;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3919; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
; QY 1 GGTCACCTGTGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60

1 GGTCACCTGTGAGAGCAGCTTCTCTGCGCGCCCTCTGCGCAAGGAGGAGACTCGAC 60  
61 AACATGTCAACCGCGCTCCCAAGACCTCTGCGCAACCCGAGGCTCTGAAGAAACCCCTCGG 120  
61 AACATGTCAACCGCGCTCCCAAGACCTCTGCGCAACCCGAGGCTCTGAAGAAACCCCTCGG 120  
121 GATGAGATCAATGCATCTTCTGCAAGAGGATATGCTGCTGATGAGGATGAGGAGGAC 180  
121 GATGAGATCAATGCATCTTCTGCAAGAGGATATGCTGCTGATGAGGATGAGGAGGAC 180  
181 ATGATCCAGCGGAGAGAGCTTAAACGAAACACCTTCCGAGGCTCAAGAAATTTAAAGATCAT 240  
181 ATGATCCAGCGGAGAGAGCTTAAACGAAACACCTTCCGAGGCTCAAGAAATTTAAAGATCAT 240  
241 GCCAGGCGCTGAAGGCAACATGCAATTTAAGTATTAAGTATTAAGTATTAAGTATTAAG 300  
241 GCCAGGCGCTGAAGGCAACATGCAATTTAAGTATTAAGTATTAAGTATTAAGTATTAAG 300  
301 CAAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTGAAACAAATATCTTTTAGC 360  
301 CAAATCCATAAGGAATACCTTGTGCTGGGCGAGATATCATTGAAACAAATATCTTTTAGC 360  
361 AGCACTAGTATTCGCCAAGCTGACTATGGCTTTGAAACACTTGGGCTACCGGATGAAATG 420  
361 AGCACTAGTATTCGCCAAGCTGACTATGGCTTTGAAACACTTGGGCTACCGGATGAAATG 420  
421 TGCTCTCAGGAGTGGCGAGAAAGCTGCCGAGGAGTAACTCTCCAGAGGAAATTAAG 480  
421 TGCTCTCAGGAGTGGCGAGAAAGCTGCCGAGGAGTAACTCTCCAGAGGAAATTAAG 480  
481 AGTTTGTGGCAGGGGCTCTGGGTCGCACTAAAGACACTCTCTGTGTCCCATCTGTG 540  
481 AGTTTGTGGCAGGGGCTCTGGGTCGCACTAAAGACACTCTCTGTGTCCCATCTGTG 540  
541 GAAAGCGCGGATTATAGGAAACATCACAATTTGATGAGCTTGTGTAAGCATCAAGAGCAG 600  
541 GAAAGCGCGGATTATAGGAAACATCACAATTTGATGAGCTTGTGTAAGCATCAAGAGCAG 600  
601 GCCAAGGAGCTTCTGGATGGCGGGTTGATATCTTCACTGATGAACTATTTTGTACT 660  
601 GCCAAGGAGCTTCTGGATGGCGGGTTGATATCTTCACTGATGAACTATTTTGTACT 660  
661 GCCAATCCAGGAGGAGCTTGTGCTGCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
661 GCCAATCCAGGAGGAGCTTGTGCTGCACTCCAAATCTTTTGGAGGAAATATGCTCCC 720  
721 CGGCTATCTTTATTTTTCAGGAGCAGATCGTTGATAAAAAGTGGGCGGACTCTTTCCGACAG 780  
721 CGGCTATCTTTATTTTTCAGGAGCAGATCGTTGATAAAAAGTGGGCGGACTCTTTCCGACAG 780  
781 ACAGGAGGAGGATTTGATCAGCGGTCTCATGGAGAACCACTCTGCAATGGATTAAT 840  
781 ACAGGAGGAGGATTTGATCAGCGGTCTCATGGAGAACCACTCTGCAATGGATTAAT 840  
841 TGCTCTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATATTTGAAATATTTGAAATATTTG 900  
841 TGCTCTTGGGTGAGCTGAGATGAGACCTTTTATTTGAAATATTTGAAATATTTGAAATATTTG 900  
901 GCCTATGTCTCTCTGTTATCCCAATGTCAGGCTCTTCCCAACACCTTTTGGTGAATGATA 960  
901 GCCTATGTCTCTCTGTTATCCCAATGTCAGGCTCTTCCCAACACCTTTTGGTGAATGATA 960  
961 ACGCTCTTATGATGGCGCAAGCACCCTAAGAGATTTTGTATGGATGGCTTGGTCAATATA 1020  
961 ACGCTCTTATGATGGCGCAAGCACCCTAAGAGATTTTGTATGGATGGCTTGGTCAATATA 1020  
1021 GTTGAGGATGCTGTGGGTCAACACCAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
1021 GTTGAGGATGCTGTGGGTCAACACCAGATCATATCAGGAAATTTGCTGAAGCTGTGAAA 1080  
1081 AATTGTAAGCTTAGAGTTTCCACCTGCTCTTTTGAAGGACATATGTTTACTGTCTGGT 1140

1081	Db	AA	TTGTTAAGCCTAGAGTTCCACCTGGCCAC	CTGCTTTTGAAGGACATATGTTACTGTC	TTGGT	1144
1141	Qy	CT	AGAGCCCTT	CAGGATTTGGACCGTACACCAACTTTGTTAACTATGGAGAGCGCTGTAAT		1200
1141	Db	CT	AGAGCCCTT	CAGGATTTGGACCGTACACCAACTTTGTTAACTATGGAGAGCGCTGTAAT		1200
1201	Qy	GT	TGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGGAAACTATGAAGAAGCCTTG			1260
1201	Db	GT	TGCAGGATCAAGGAAGTTTGTCTAAACTCATCATGCGAGGAAACTATGAAGAAGCCTTG			1260
1261	Qy	TG	TGTTGCCAAAGTGCAGGTGCGAAATGGGAGCCACAGTGTGGAATGTCACATCGATGAT			1320
1261	Db	TG	TGTTGCCAAAGTGCAGGTGCGAAATGGGAGCCACAGTGTGGAATGTCACATCGATGAT			1320
1321	Qy	GG	CATGCTAGATGTTCCAAGTGCATGACCGAATTTTGCACACTTAAATGCTTCCGAGCCA			1380
1321	Db	GG	CATGCTAGATGTTCCAAGTGCATGACCGAATTTTGCACACTTAAATGCTTCCGAGCCA			1380
1381	Qy	GAC	ATCGCAAAAGTATCTTTTGTGATCGATCTCTCCAAATTTGCTGTGATTTGAAGCTGGG			1440
1381	Db	GAC	ATCGCAAAAGTATCTTTTGTGATCGATCTCTCCAAATTTGCTGTGATTTGAAGCTGGG			1440
1441	Qy	TT	AAAGTGTGCCAAAGGAGTGCATTTGTCATATAGCATTAAGTCTGAGGAGGAGAGGAC			1500
1441	Db	TT	AAAGTGTGCCAAAGGAGTGCATTTGTCATATAGCATTAAGTCTGAGGAGGAGAGGAC			1500
1501	Qy	GAC	TTCTTGGAGAGGCCAGGAAGATTAAAGATGATGGAGCTGTATGTGGTGGTCA	TGGCT		1560
1501	Db	GAC	TTCTTGGAGAGGCCAGGAAGATTAAAGATGATGGAGCTGTATGTGGTGGTCA	TGGCT		1560
1561	Qy	TT	TGATGAAGAAGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCA	CCCGGGCC		1620
1561	Db	TT	TGATGAAGAAGACAGGCAACAGAAACAGACACAAAAATCAGAGTGTGCA	CCCGGGCC		1620
1621	Qy	TAC	CATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATATTTTGTGACCCCTAAT			1680
1621	Db	TAC	CATCTGCTTGTGAAAACCTGGGCTTTAATCCAAATGACATATTTTGTGACCCCTAAT			1680
1681	Qy	AT	CTTAACCATTTGGGACTGGAAATGGAGGAAACAACTTGTATGCCATTAATTTTATCCAT			1740
1681	Db	AT	CTTAACCATTTGGGACTGGAAATGGAGGAAACAACTTGTATGCCATTAATTTTATCCAT			1740
1741	Qy	GCA	CAAAAGTCATTAAGAAACATTA	CTGGAGCCAGATAAGTGGAGTCTTTTCCAAC		1800
1741	Db	GCA	CAAAAGTCATTAAGAAACATTA	CTGGAGCCAGATAAGTGGAGTCTTTTCCAAC		1800
1801	Qy	TT	GTCTTCTCTTCCGAGGAAATGGAAGCCCACTCGAAGCAATGCA	TGGGTTTTCCTT		1860
1801	Db	TT	GTCTTCTCTTCCGAGGAAATGGAAGCCCACTCGAAGCAATGCA	TGGGTTTTCCTT		1860
1861	Qy	TAC	CATGCAATCAAGTCTGGCATGCGACATGGAGATAGTGAATGCTGGAACCTCCCTGTG			1920
1861	Db	TAC	CATGCAATCAAGTCTGGCATGCGACATGGAGATAGTGAATGCTGGAACCTCCCTGTG			1920
1921	Qy	TAT	GATGATATCCAATAAGGAACCTTCTG	CAGCTCTGTGAAGATCTCATCTCGAAATAAAGAC		1980
1921	Db	TAT	GATGATATCCAATAAGGAACCTTCTG	CAGCTCTGTGAAGATCTCATCTCGAAATAAAGAC		1980
1981	Qy	CCT	GAGGCCACTGAGAGACTCTTACGTTATGCCAGACTC	AGGCCACAGAGGGAAGAAA		2040
1981	Db	CCT	GAGGCCACTGAGAGACTCTTACGTTATGCCAGACTC	AGGCCACAGAGGGAAGAAA		2040
2041	Qy	GT	CATTGACACTGATGATGGAGAAATGGCCCTGTGCAAGAACCCCTTGAGTATG	CGCCCTT		2100
2041	Db	GT	CATTGACACTGATGATGGAGAAATGGCCCTGTGCAAGAACCCCTTGAGTATG	CGCCCTT		2100
2101	Qy	GT	GAGGGCATTGAAAACATATTAATTTGAGGATACTGAGGAAGCCAGTGTAAACCAAAA			2160
2101	Db	GT	GAGGGCATTGAAAACATATTAATTTGAGGATACTGAGGAAGCCAGTGTAAACCAAAA			2160
2161	Qy	AA	ATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGATGATGAATGAATGAAATGTT			2220
2161	Db	AA	ATATCCCGACCTCTCAATATATTAATTTGAAGGACCCCTGATGATGAATGAATGAAATGTT			2220

Qy	2221	GGTGAATCTTTTGGAGCTGGAATAATGTTTCTACTCAAGTTATATAAGTCAGCCCGGGTT	2280
Db	2221	GGTGATCTTTTGGAGCTGGAATAATGTTTCTACTCAAGTTATATAAGTCAGCCCGGGTT	2280
Qy	2281	ATGAAGAAGGCTGTTCGGCCACCTTATCCCTTTTCATGCAAAAAGAGAAGAACCCAGA	2340
Db	2281	ATGAAGAAGGCTGTTCGGCCACCTTATCCCTTTTCATGCAAAAAGAGAAGAACCCAGA	2340
Qy	2341	GTGCTTAAACGGCACAGTAGAAGAAGAGGACCCCTTACCAGGGCACCATCGTGTGGCCACT	2400
Db	2341	GTGCTTAAACGGCACAGTAGAAGAAGAGGACCCCTTACCAGGGCACCATCGTGTGGCCACT	2400
Qy	2401	GTTAAAGGCACGTGCACGACATAGGCAAGAACATAGTTTGGAGTAGTCTCTGGCTGCAT	2460
Db	2401	GTTAAAGGCACGTGCACGACATAGGCAAGAACATAGTTTGGAGTAGTCTCTGGCTGCAT	2460
Qy	2461	AAATTCGGAGTTATGATTTTAGGAGTCATGACTCCCATGTGATAAGTACTGAAAGCTGCT	2520
Db	2461	AAATTCGGAGTTATGATTTTAGGAGTCATGACTCCCATGTGATAAGTACTGAAAGCTGCT	2520
Qy	2521	CTTGACCAAAAGCAGATATAATTTGGCCCTGTCCAGGACTCATCACTCTCTCCCTGGATGAA	2580
Db	2521	CTTGACCAAAAGCAGATATAATTTGGCCCTGTCCAGGACTCATCACTCTCTCCCTGGATGAA	2580
Qy	2581	ATGATTTTGTTCGCAAGGAAATGGAGAGATTAGCTTATAAGGATTCATTTGTTGATTGGA	2640
Db	2581	ATGATTTTGTTCGCAAGGAAATGGAGAGATTAGCTTATAAGGATTCATTTGTTGATTGGA	2640
Qy	2641	GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAAGCTCCGAGATACAGTGCACT	2700
Db	2641	GGAGCAACCACTTCAAAAACCCACACAGCAGTTAAATAAGCTCCGAGATACAGTGCACT	2700
Qy	2701	GTAATCCATGCTCTGGACGGCTCCAAAGAGTGTGGTGTCTCCAGCTGTTCCAGCTGTTAGATGAA	2760
Db	2701	GTAATCCATGCTCTGGACGGCTCCAAAGAGTGTGGTGTCTCCAGCTGTTCCAGCTGTTAGATGAA	2760
Qy	2761	AACTCTAAAGGATGAATACTTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGAC	2820
Db	2761	AACTCTAAAGGATGAATACTTTTGAGGAAATCATGGAAGAATATGAAGATATTAGACAGGAC	2820
Qy	2821	CATTATGAGTCTCTCAAGAGAGGAGATACTTACCCCTTAAGTCAAGCCAGCAAAAAGTGCT	2880
Db	2821	CATTATGAGTCTCTCAAGAGAGGAGATACTTACCCCTTAAGTCAAGCCAGCAAAAAGTGCT	2880
Qy	2881	TTCCAAATGATTTGGTGTCTGAACCTCAACCAAGTGAAGCCCAAGCTTATATGGAGCCAG	2940
Db	2881	TTCCAAATGATTTGGTGTCTGAACCTCAACCAAGTGAAGCCCAAGCTTATATGGAGCCAG	2940
Qy	2941	GTCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAGGCTTTCTTTT	3000
Db	2941	GTCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTTGACTGGAGGCTTTCTTTT	3000
Qy	3001	GATGTCGAGCTCCGGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTTAAACGAC	3060
Db	3001	GATGTCGAGCTCCGGGGCAAGTACCGAATCGAGGCTTCCCAAGATATTTTAAACGAC	3060
Qy	3061	AAACAGTAGTGGAGGCGCCAGGAGGCTACGATGATGCCCAAAATATGCTGAAACACA	3120
Db	3061	AAACAGTAGTGGAGGCGCCAGGAGGCTACGATGATGCCCAAAATATGCTGAAACACA	3120
Qy	3121	CTGATTAGTCAAAAAGAACTCCGGGCCCGGGGTGTGTTGGGTCTTGGCCAGCACAGAGT	3180
Db	3121	CTGATTAGTCAAAAAGAACTCCGGGCCCGGGGTGTGTTGGGTCTTGGCCAGCACAGAGT	3180
Qy	3181	ATCCAGACGACATTCACCTGTACGACAGGCTGTGTGCCCAAGCTGACAGGCCCATAC	3240
Db	3181	ATCCAGACGACATTCACCTGTACGACAGGCTGTGTGCCCAAGCTGACAGGCCCATAC	3240
Qy	3241	GCCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCGCAGCAGGAGCCATAC	3300
Db	3241	GCCACTTTCTATGGGTTAAGGCAACAGGCTGAGAGGACTCTGCGCAGCAGGAGCCATAC	3300









Qy 3184 CAAGACGACATTCACCTGTAGCGAGAGCTGTGTGCCCGAGGCTGCAGAGCCCATAGCC 3243  
Db 3121 CAAGACGACATTCACCTGTAGCGAGAGCTGTGTGCCCGAGGCTGCAGAGCCCATAGCC 3180  
Qy 3244 ACTTTCTATGGTTAAGCAACAGGCTGAGAAGACTCTGCGAGCAGCGAGCCATATCTAC 3303  
Db 3181 ACTTTCTATGGTTAAGCAACAGGCTGAGAAGACTCTGCGAGCAGCGAGCCATATCTAC 3240  
Qy 3304 TGCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGCAATCTCGCATCCGTGACTACCTGGGGCTGTTT 3363  
Db 3241 TGCCTCTCAGACTTCATCGCTCCCTTGCATTTCTGCAATCTCGCATCCGTGACTACCTGGGGCTGTTT 3300  
Qy 3364 GCCGTGCTCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAGAC 3423  
Db 3301 GCCGTGCTCTGCTTTGGGGTAGAAGAGCTGAGCAAGGCTATGAGGATGATGGTGAAGAC 3360  
Qy 3424 TACAGCAGCATATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAGAGAG 3483  
Db 3361 TACAGCAGCATATGGTCAAGCGCTGGGGACCGGCTGGCAGAGGCTTTTGCAGAGAG 3420  
Qy 3484 CTCCATGAAGAGTTCCGCGAGAACTGTGGGCTTACTGTGCGAGTGAGCAGCTGGACGTC 3543  
Db 3421 CTCCATGAAGAGTTCCGCGAGAACTGTGGGCTTACTGTGCGAGTGAGCAGCTGGACGTC 3480  
Qy 3544 GCAGACCTGCGAAGGTTGCGGTACAAAGGCACTCCGCGGCTCCTGGCTACCCCGAGCCAG 3603  
Db 3481 GCAGACCTGCGAAGGTTGCGGTACAAAGGCACTCCGCGGCTCCTGGCTACCCCGAGCCAG 3540  
Qy 3604 CCGACACACCGAGAACTCACCATGTGGAGACTCGCAGACATCGCAGCATCGAGCATCTACAGGC 3663  
Db 3541 CCGACACACCGAGAACTCACCATGTGGAGACTCGCAGACATCGCAGCATCGAGCATCTACAGGC 3600  
Qy 3664 ATTAGTTAAACAGATCATTAAGCAATGCGACCTGCTTCCAGCATCTCAGGCTCTACTTC 3723  
Db 3601 ATTAGTTAAACAGATCATTAAGCAATGCGACCTGCTTCCAGCATCTCAGGCTCTACTTC 3660  
Qy 3724 TCCAAATTTGAAGTCCAAATATTTTGTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGGAT 3783  
Db 3661 TCCAAATTTGAAGTCCAAATATTTTGTGTGGGAGAGATTTCCAAAGGATCAGGTTGAGGAT 3720  
Qy 3784 TATGCAATTTGAAGAGAAACATATCTGTGCTGAGGTTGAGAAATGCTTTGGACCCATTTTG 3843  
Db 3721 TATGCAATTTGAAGAGAAACATATCTGTGCTGAGGTTGAGAAATGCTTTGGACCCATTTTG 3780  
Qy 3844 GGATATGATACAGACTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGCTTTTATCTTGATGATCCTCA 3903  
Db 3781 GGATATGATACAGACTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGCTTTTATCTTGATGATCCTCA 3840  
Qy 3904 AGGAATACAACTTAG 3919  
Db 3841 AGGAATACAACTTAG 3856

RESULT 3  
US-09-796-692-8171  
; Sequence 8171, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Mannion, Jane  
; APPLICANT: Algate, Paul A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8171  
; LENGTH: 420  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-8171

Query Match 8.1%; Score 318; DB 3; Length 420;  
Best Local Similarity 99.5%; Pred. No. 3.5e-153;  
Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2749 CTGTTAGATGAAATCTAAAGGATGAATCTTTTGGAGAAATCATGGAAGAAATATGAAGAT 2808  
Db 1 CTGTTAGATGAAATCTAAAGGATGAATCTTTTGGAGAAATCATGGAAGAAATATGAAGAT 60  
Qy 2809 ATTAGACAGACCATTAATGATGCTCTCAAGAGAGGAGATCTTACCTTTAAGTCAAGCC 2868  
Db 61 ATTAGACAGGCGCAATTAATGATGCTCTCAAGAGAGGAGATCTTACCTTTAAGTCAAGCC 120  
Qy 2869 AGAAAAAGTGGTTTCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 2928  
Db 121 AGAAAAAGTGGTTTCCAAATGGAATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 180  
Qy 2929 ATTGGGACCCAGGCTCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTGACTGG 2988  
Db 181 ATTGGGACCCAGGCTCTTTGAAGACTATGACCTGCAGAAAGCTGGTGGACTACATTGACTGG 240  
Qy 2989 AAGCTTTCTTTGATGCTGCGAGCTCCGGGGCAAGTACCGGAATCGAGGCTTCCCAAG 3048  
Db 241 AAGCTTTCTTTGATGCTGCGAGCTCCGGGGCAAGTACCGGAATCGAGGCTTCCCAAG 300  
Qy 3049 ATATTTAAGCAAAACAGTAGGTGGAGAGCCAGGAGGCTTACGATGATGCCCAAT 3108  
Db 301 ATATTTAAGCAAAACAGTAGGTGGAGAGCCAGGAGGCTTACGATGATGCCCAAT 360  
Qy 3109 ATGCTGAACACACTGATTAGTCAAAAGAAATCTCCGGGCGCGGGGTGTGGTCTTCTGG 3168  
Db 361 ATGCTGAACACACTGATTAGTCAAAAGAAATCTCCGGGCGCGGGGTGTGGTCTTCTGG 420

RESULT 4  
US-10-040-862-8171  
; Sequence 8171, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014059-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862

```
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8171
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8171

Query Match      8.1%; Score 318; DB 5; Length 420;
Best Local Similarity 99.5%; Pred. No. 3.5e-153;
Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2749 CTGTTAGATCAAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAGATATGAAGAT 2808
DB 1 CTGTTAGATCAAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAGATATGAAGAT 60

QY 2809 ATTAGACAGGACCAATATGATCTCTCAAGGAGGAGATCTTACCCCTTAAGTCAAGCC 2868
DB 61 ATTAGACAGGACCAATATGATCTCTCAAGGAGGAGATCTTACCCCTTAAGTCAAGCC 120

QY 2869 AGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCACCAGTGAAGCCACGTTT 2928
DB 121 AGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCACCAGTGAAGCCACGTTT 180

QY 2929 ATTGGGACCCAGGCTCTTTGAGACTATGACTTCAGAGCTGGTGGACTACATTGACTGG 2988
DB 181 ATTGGGACCCAGGCTCTTTGAGACTATGACTTCAGAGCTGGTGGACTACATTGACTGG 240

QY 2989 AAGCTTTTCTTTGATGTCTGACGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAG 3048
DB 241 AAGCTTTTCTTTGATGTCTGACGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAG 300

QY 3049 ATATTAAACGACAAAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 3108
DB 301 ATATTAAACGACAAAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 360

QY 3109 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCCCGGGTGTGGTGGTCTTGG 3168
DB 361 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCCCGGGTGTGGTGGTCTTGG 420.
```

## RESULT 5

```
US-10-057-475B-8171
; Sequence 8171, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8171
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-8171
```

```
Query Match      8.1%; Score 318; DB 6; Length 420;
Best Local Similarity 99.5%; Pred. No. 3.5e-153;
Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2749 CTGTTAGATCAAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAGATATGAAGAT 2808
DB 1 CTGTTAGATCAAAATCTAAAGGATGAATCTTTGAGGAAATCATGGAAGATATGAAGAT 60

QY 2809 ATTAGACAGGACCAATATGATCTCTCAAGGAGGAGATCTTACCCCTTAAGTCAAGCC 2868
DB 61 ATTAGACAGGACCAATATGATCTCTCAAGGAGGAGATCTTACCCCTTAAGTCAAGCC 120

QY 2869 AGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCACCAGTGAAGCCACGTTT 2928
DB 121 AGAAAAAGTGGTTTCCAAATGGATGGCTGTCTGAACCTCACCAGTGAAGCCACGTTT 180

QY 2929 ATTGGGACCCAGGCTCTTTGAGACTATGACTTCAGAGCTGGTGGACTACATTGACTGG 2988
DB 181 ATTGGGACCCAGGCTCTTTGAGACTATGACTTCAGAGCTGGTGGACTACATTGACTGG 240

QY 2989 AAGCTTTTCTTTGATGTCTGACGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAG 3048
DB 241 AAGCTTTTCTTTGATGTCTGACGCTCGGGGCAAGTACCCGAATCGAGGCTTCCCAAG 300

QY 3049 ATATTAAACGACAAAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 3108
DB 301 ATATTAAACGACAAAACAGTAGGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 360

QY 3109 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCCCGGGTGTGGTGGTCTTGG 3168
```

Db 361 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCGGGGTGTGGTTGGTTCTGG 420

RESULT 6

US-10-154-884B-8171

; Sequence 8171, Application US/10154884B

; Publication No. US2004000561A1

GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; APPLICANT: Retter, Marc W.

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

; TITLE OF INVENTION: Hematological Malignancies

; FILE REFERENCE: 014058-013521US

; CURRENT FILING DATE: 2002-05-23

; PRIOR APPLICATION NUMBER: US/10/154,884B

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: US 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: US 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: US 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: US 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: US 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/222,903

; PRIOR FILING DATE: 2000-08-03

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 11290

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8171

; LENGTH: 420

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-154-884B-8171

Query Match 8.1%; Score 318; DB 6; Length 420;

Best Local Similarity 99.5%; Pred. No. 3.5e-153;

Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2749 CTGTTAGATGAAATCTTAAGATGAATACCTTTGAGGAAATCATGGAAGAATATGAAGAT 2808

Db 1 CTGTTAGATGAAATCTTAAGATGAATACCTTTGAGGAAATCATGGAAGAATATGAAGAT 60

QY 2809 ATTACACAGGACCATATGAGTCTCTCAAGGAGAGAGATACCTTACCCTTAAGTCAAGCC 2868

Db 61 ATTACACAGGACCATATGAGTCTCTCAAGGAGAGAGATACCTTACCCTTAAGTCAAGCC 120

QY 2869 AGAAAAAGTGGTTTCCAAATGATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 2928

Db 121 AGAAAAAGTGGTTTCCAAATGATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 180

QY 2929 ATTGGGACCCAGGCTTTTGAAGACTATGACCTGCGAGAGAGCTGGTGGACTACATTGACTGG 2988

Db 181 ATTGGGACCCAGGCTTTTGAAGACTATGACCTGCGAGAGAGCTGGTGGACTACATTGACTGG 240

QY 2989 AAGCCTTTCTTTGATGCTCTGGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTTCCCAAG 3048

Db 241 AAGCCTTTCTTTGATGCTCTGGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTTCCCAAG 300

QY 3049 ATATTTAACGACAAACAGTAGTGGAGAGGCCAGGAAGGTTACGATGATGCCCAACAAT 3108

Db 301 ATATTTAACGACAAACAGTAGTGGAGAGCCAGGAAGGTTACGATGATGCCCAACAAT 360

QY 3109 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCGGGGTGTGGTTGGTTCTGG 3168

Db 361 ATGCTGAACACACTGATTAGTCAAAAGAAACTCCGGGCCGGGGTGTGGTTGGTTCTGG 420

RESULT 7

US-10-764-324-8171

; Sequence 8171, Application US/10764324

; Publication No. US2004017539A1

GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; APPLICANT: Retter, Marc

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

; TITLE OF INVENTION: Hematological Malignancies

; FILE REFERENCE: 014058-013520US

; CURRENT FILING DATE: 2004-01-23

; PRIOR APPLICATION NUMBER: US/10/764,324

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: US 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: US 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: US 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: US 60/218,950

; PRIOR FILING DATE: 2000-07-14

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 10467

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8171

; LENGTH: 420

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-764-324-8171

Query Match 8.1%; Score 318; DB 8; Length 420;

Best Local Similarity 99.5%; Pred. No. 3.5e-153;

Matches 418; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2749 CTGTTAGATGAAATCTTAAGATGAATACCTTTGAGGAAATCATGGAAGAATATGAAGAT 2808

Db 1 CTGTTAGATGAAATCTTAAGATGAATACCTTTGAGGAAATCATGGAAGAATATGAAGAT 60

QY 2809 ATTACACAGGACCATATGAGTCTCTCAAGGAGAGAGATACCTTACCCTTAAGTCAAGCC 2868

Db 61 ATTACACAGGACCATATGAGTCTCTCAAGGAGAGAGATACCTTACCCTTAAGTCAAGCC 120

QY 2869 AGAAAAAGTGGTTTCCAAATGATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 2928

Db 121 AGAAAAAGTGGTTTCCAAATGATGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTT 180

QY 2929 ATTGGGACCCAGGCTTTTGAAGACTATGACCTGCGAGAGAGCTGGTGGACTACATTGACTGG 2988

Db 181 ATTGGGACCCAGGCTTTTGAAGACTATGACCTGCGAGAGAGCTGGTGGACTACATTGACTGG 240

QY 2989 AAGCCTTTCTTTGATGCTCTGGCAGCTCCGGGGCAAGTACCCGAAATCGAGGCTTTCCCAAG 3048

Db 241 AAGCTTTCTTTGATGTCGACGCTCGGGGCAAGTACCCGAATCGAGGCTTTCCTCAAG 300  
Qy 3049 ATATTAAACGAAAAACAGTAGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 3108  
Db 301 ATATTAAACGAAAAACAGTAGTGGAGAGCCAGGAGGTCTACGATGATGCCCAAT 360  
Qy 3109 ATGCTGAACACACTGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTTGGGTTCTGG 3168  
Db 361 ATGCTGAACACACTGATTAGTCAAAAGAACTCCGGGCCCGGGGTGTTGGGTTCTGG 420

## RESULT 8

US-10-040-739-642  
; Sequence 642, Application US/10040739  
; Publication No. US20020173635A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John  
; LaVallie, Edward  
; Racie, Lisa  
; Merberg, David  
; Treacy, Maurice  
; Spaulding, Vikki  
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS  
; NUMBER OF SEQUENCES: 1519  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/040,739  
; FILING DATE: 07-Jan-2002  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/036,520  
; FILING DATE: 03-JUN-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 642:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 305 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 642:  
US-10-040-739-642

Query Match 5.7%; Score 222; DB 5; Length 305;  
Best Local Similarity 99.6%; Pred. No. 1.9e-103;  
Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2751 GTTAGTCAAAATCTAAAGGATGATCTTTGAGGAATCATGGAGNATATGAGATAT 2810  
Db 22 GTTAGTCAAAATCTAAAGGATGATCTTTGAGGAATCATGGAGNATATGAGATAT 81  
Qy 2811 TAGACAGGACCATATATAGTCTCTCAAGGAGGAGATACCTTACCTTAAGTCAAGCCAG 2870  
Db 82 TAGACAGGACCATATATAGTCTCTCAAGGAGGAGATACCTTACCTTAACTCAAGCCAG 141  
Qy 2871 AAAAAAGTGGTTTCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACGCTTTAT 2930

Db 142 AAAAAAGTGGTTTCCAAATGGATTGGCTGTCTGAACCTCACCCAGTGAAGCCACGTTTAT 201  
Qy 2931 TGGGACCACAGGTCTTTGAAGACTATGACCTGCGAGAGCTGCTGGACTACATTGACTCGAA 2990  
Db 202 TGGGACCACAGGTCTTTGAAGACTATGACCTGCGAGAGCTGCTGGACTACATTGACTCGAA 261  
Qy 2991 GCCTTTCTTTGATGTCGCGAGCTCCGGGGCAA 3023  
Db 262 GCCTTTCTTTGATGTCGCGAGCTCCGGGGCAA 294

## RESULT 9

US-10-027-632-121776  
; Sequence 121776, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; POLYMORPHISMS IN THE HUMAN GENOME  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 121776  
; LENGTH: 901  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(901)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-121776

Query Match 3.2%; Score 124; DB 5; Length 901;  
Best Local Similarity 100.0%; Pred. No. 1.2e-52;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1893 GATAGTGAATGCTGGAAACCTCCCTGTGTATGATATCCATTAAGGAACCTTCTGCAGCT 1952  
Db 359 GATAGTGAATGCTGGAAACCTCCCTGTGTATGATATCCATTAAGGAACCTTCTGCAGCT 418  
Qy 1953 CTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGAGGCTCTTACGTTATGC 2012  
Db 419 CTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGAGGCTCTTACGTTATGC 478  
Qy 2013 CCAG 2016  
Db 479 CCAG 482

## RESULT 10

US-10-027-632-121776  
; Sequence 121776, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121776
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(901)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-121776

Query Match          3.2%; Score 124; DB 6; Length 901;
Best Local Similarity 100.0%; Pred. No. 1.2e-52;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1893 GATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATCCATAGGAACCTTCTGCAGCT 1952
Db      |||||||
QY 1953 CTGTGAAGATCTCATCTGGAATAAAGACCTTGAGGCCACTGAGAGCTCTTACGTTATGC 2012
Db      |||||||
QY 2013 CCAG 2016
Db      |||||
QY 479 CCAG 482
Db      |||||

RESULT 11
US-09-925-065A-727226/c
; Sequence 727226, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 697175
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-697175

Query Match          2.0%; Score 80; DB 4; Length 616;
Best Local Similarity 100.0%; Pred. No. 7.6e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2739 GTGTTCCAGCTGTTAGATGAAATCTAAAGAGTAATCTTGAAGAAATCATGGAAGA 2798
Db      |||||||
QY 2799 ATATGAAGATATTAGACAGG 2818
Db      |||||||
QY 74 ATATGAAGATATTAGACAGG 55
Db      |||||||

RESULT 13
US-10-027-632-56722/c
; Sequence 56722, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```



```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56722
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-56722

Query Match      1.8%; Score 69; DB 5; Length 493;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      862 ATGAGACCTTTTATTGGAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 921
Db      132 ATGAGACCTTTTATTGGAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 73

Qy      922 AATGCAGGT 930
Db      72 AATGCAGGT 64

RESULT 15
US-10-027-632-314465/c
; Sequence 314465, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 314465
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-314465

Query Match      1.8%; Score 69; DB 5; Length 493;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      862 ATGAGACCTTTTATTGGAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 921
Db      132 ATGAGACCTTTTATTGGAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 73

Qy      922 AATGCAGGT 930
Db      72 AATGCAGGT 64

RESULT 14
US-10-027-632-314464/c
; Sequence 314464, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 314464
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-314464
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56722
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-56722

Query Match
Best Local Similarity 100.0%; Score 69; DB 6; Length 493;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 862 ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTTATCCC 921
Db 132 ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTTATCCC 73

QY 922 AATGCAGGT 930
Db 72 AATGCAGGT 64

RESULT 18
US-10-027-632-314465/c
; Sequence 314465, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 314465
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-314465

Query Match
Best Local Similarity 100.0%; Score 69; DB 6; Length 493;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 862 ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTTATCCC 921
Db 132 ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTTATCCC 73

QY 922 AATGCAGGT 930
Db 72 AATGCAGGT 64

RESULT 19
US-10-027-632-3760/c
; Sequence 3760, Application US/10027632
```

```
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3760
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(627)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-3760

Query Match      1.8%   Score 69;   DB 5;   Length 627;
Best Local Similarity 100.0%;   Pred. No. 3.8e-24;
Matches 69;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      862  ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 921
      |||||||
DB      271  ATGAGACCTTTTATTGAAATAATTGGAATAATGTACACAGCCTATGCTCTGTATATCCC 212

QY      922  AATGCAGGT 930
      |||||||
DB      211  AATGCAGGT 203

RESULT 20
US-10-027-632-3760/c
; Sequence 3760, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3760
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 460077
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-460077

Query Match      1.0%   Score 39;   DB 4;   Length 611;
Best Local Similarity 100.0%;   Pred. No. 1.4e-08;
Matches 39;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      989  AGGATTTTGTATGATGGCTTGTCAATATAGTTGAG 1027
      |||||||
DB      39  AGGATTTTGTATGATGGCTTGTCAATATAGTTGAG 1

RESULT 21
US-09-925-065A-460077/c
; Sequence 460077, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 460077
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-460077

Query Match      1.0%   Score 39;   DB 4;   Length 611;
Best Local Similarity 100.0%;   Pred. No. 1.4e-08;
Matches 39;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      989  AGGATTTTGTATGATGGCTTGTCAATATAGTTGAG 1027
      |||||||
DB      39  AGGATTTTGTATGATGGCTTGTCAATATAGTTGAG 1

RESULT 22
US-09-925-065A-39453/c
; Sequence 39453, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```

; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39453  
; LENGTH: 588  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-39453

Query Match 0.8%; Score 31; DB 4; Length 588;  
Best Local Similarity 100.0%; Pred. No. 0.00019;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3889 TCTTGATGATCCTCAAGGAATACAACTAG 3919  
Db 570 TCTTGATGATCCTCAAGGAATACAACTAG 540

RESULT 23  
US-10-607-712-45/c  
; Sequence 45, Application US/10607712  
; Publication No. US20040073018A1  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002005  
; CURRENT APPLICATION NUMBER: US/10/607,712  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 08/980,326  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/031,964  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: 60/050,310  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 28  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-607-712-45

Query Match 0.7%; Score 28; DB 7; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0063;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2468 GAGTTATTGATTAGGATGATGACTCC 2495  
Db 28 GAGTTATTGATTAGGATGATGACTCC 1

RESULT 24  
US-10-607-712-35  
; Sequence 35, Application US/10607712  
; Publication No. US20040073018A1  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima

; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002005  
; CURRENT APPLICATION NUMBER: US/10/607,712  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 08/980,326  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/031,964  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: 60/050,310  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-607-712-35

Query Match 0.7%; Score 26; DB 7; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 TTCCGAGGTCAGGAATTTAAAGATCA 239  
Db 1 TTCCGAGGTCAGGAATTTAAAGATCA 26

RESULT 25  
US-10-607-712-48  
; Sequence 48, Application US/10607712  
; Publication No. US20040073018A1  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; FILE REFERENCE: 50004/002005  
; CURRENT APPLICATION NUMBER: US/10/607,712  
; PRIOR FILING DATE: 2003-06-27  
; PRIOR APPLICATION NUMBER: 08/980,326  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: 60/031,964  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: 60/050,310  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 48  
; LENGTH: 26  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-607-712-48

Query Match 0.7%; Score 26; DB 7; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2422 ATAGCGAAGAACATAGTTGGAGTAGT 2447  
Db 1 ATAGCGAAGAACATAGTTGGAGTAGT 26

RESULT 26  
US-10-607-712-53

```

: Sequence 53, Application US/10607712
: Publication No. US20040073018A1
: GENERAL INFORMATION:
: APPLICANT: Gravel, Roy A.
: APPLICANT: Rozen, Rima
: APPLICANT: LeClerc, Daniel
: APPLICANT: Goyette, Philippe
: APPLICANT: Campeau, Eric
: TITLE OF INVENTION: HUMAN METHIONIN
: TITLE OF INVENTION: METHODS FOR EV
: TITLE OF INVENTION: CARDIOVASCULAR
: FILE REFERENCE: 50004/002005
: CURRENT APPLICATION NUMBER: US/10/6
: CURRENT FILING DATE: 2003-06-27
: PRIOR APPLICATION NUMBER: 08/980,32
: PRIOR FILING DATE: 1997-11-26
: PRIOR APPLICATION NUMBER: 60/031,96
: PRIOR FILING DATE: 1996-11-27
: PRIOR APPLICATION NUMBER: 60/050,31
: PRIOR FILING DATE: 1997-06-20
: NUMBER OF SEQ ID NOS: 76
: SOFTWARE: FastSeq for Windows Versi
: SEQ ID NO 53
: LENGTH: 26
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-607-712-53

```

Query Match 0.7%; Score 26; DB 7; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2790 CATGGAAGAATATGAAGATATTAGAC 2815  
|||  
Db 1 CATGGAAGAATATGAAGATATTAGAC 26

```

RESULT 27
US-10-607-712-57/c
; Sequence 57, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 26
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-57

```

Qy 2496 ATGTGATAAGATACTGAAAGCTGCTC 2521

Db 26 ATGTGATAAGATACTGAAAGCTGCTC 1

```

RESULT 28
US-10-607-712-63/c
; Sequence 63, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 26
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-63

```

Query Match 0.7%; Score 26; DB 7; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 26; Conservative 0; Mismatches 0; Indels

Qy 1177 GTTAAcATTGGAGAGCGCTGTAATGT 1202  
|||  
Db 26 GTTAAcATTGGAGAGCGCTGTAATGT 1

```

RESULT 29
US-09-925-065A-602656
; Sequence 602656, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 602656
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-602656

```

Query Match	0.7%;	Score 26;	DB 4;	Length 559;
Best Local Similarity	100.0%;	Pred. No. 0.074;		

Query Match	0.7%;	Score 26;	DB 4;	Length 559;
Best Local Similarity	100.0%;	Pred. No. 0.074;		

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3864 TTTTCTTTTCTTTTGGCTTTTAT 3889  
|||  
Db 21 TTTTCTTTTCTTTTGGCTTTTAT 46

```

RESULT 30
US-09-925-065A-271845
; Sequence 271845, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271845
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271845

```

Query Match 0.7%; Score 26; DB 4; Length 595;  
Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 26; Conservative 0; Mismatches 0; Indels

Qy 3863 TTTTTTTTTTTTTTTTGCCTTTT 3888  
|||  
Db 378 TTTTTTTTTTTTTTTTGCCTTTT 403

```

RESULT 31
US-10-437-963-1741/c
; Sequence 1741, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 1741
; LENGTH: 649
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101579C.1
US-10-437-963-1741

```

Query Match 0.7%; Score 26; DB 7; Length 649;

Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 26: Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;

	3862	CCTTTTCTTTTTTTTTTGCCCTTTT	.3887
Qy			
Db	305	CCTTTTCTTTTTTTTTTGCCCTTTT	280

```

RESULT 32
US-09-925-065A-702090
; Sequence 702090, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 702090
; LENGTH: 1374
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-702090

```

Query Match 0.7%; Score 26; DB 4; Length 1374;  
Best Local Similarity 100.0%; Pred. No. 0.075;  
Matches 26; Conservative 0; Mismatches 0; Indels

**Qy** 3862 CTTTCTTTTTTTTTTTTTTGCCCTTTTTT 3887  
|||  
**Db** 204 CTTTCTTTTTTTTTTTTTTGCCCTTTTTT 229

```

RESULT 33
US-10-607-712-39
; Sequence 39, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 25
; TYPE: DNA

```



```
; ORGANISM: Homo sapiens
US-10-607-712-39

Query Match      0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2233 GGAGCTGGAAATGTTTCTACCTC 2257
Db 1 GGAGCTGGAAATGTTTCTACCTC 25

RESULT 34
US-10-607-712-41/c
; Sequence 41, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: Leclerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-41

Query Match      0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2168 CCCGACCTCTCAATATTAATGAAG 2192
Db 25 CCCGACCTCTCAATATTAATGAAG 1

RESULT 35
US-10-607-712-49/c
; Sequence 49, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: Leclerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
```

```
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-49

Query Match      0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2737 GTGTGTTCCAGCTGTAGATGAA 2761
Db 25 GTGTGTTCCAGCTGTAGATGAA 1

RESULT 36
US-10-607-712-58/c
; Sequence 58, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: Leclerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-58

Query Match      0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1696 ACTGGAATGGAGGACACAACTTGT 1720
Db 25 ACTGGAATGGAGGACACAACTTGT 1

RESULT 37
US-10-607-712-59
; Sequence 59, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: Leclerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
```

```
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-59

Query Match          0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1188 AGAGCGCTGTAATGTTGCAGGATCA 1212
Db 1 AGAGCGCTGTAATGTTGCAGGATCA 25

RESULT 38
US-10-607-712-60/c
; Sequence 60, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-60

Query Match          0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2096 CCCTTGTGAAGGCGATTGAAAAACA 2120
Db 25 CCCTTGTGAAGGCGATTGAAAAACA 1

RESULT 39
US-10-607-712-61
; Sequence 61, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
```

```
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-61

Query Match          0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1527 TAAAAAGTATGAGCTGCTATGGTG 1551
Db 1 TAAAAAGTATGAGCTGCTATGGTG 25

RESULT 40
US-10-607-712-62/c
; Sequence 62, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-62

Query Match          0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1117 GAAGGACATATGTTACTGCTCTGGTC 1141
Db 25 GAAGGACATATGTTACTGCTCTGGTC 1

RESULT 41
US-10-607-712-65/c
; Sequence 65, Application US/10607712
; Publication No. US20040073018A1
; GENERAL INFORMATION:
; APPLICANT: Gravel, Roy A.
; APPLICANT: Rozen, Rima
```

```
; APPLICANT: LeClerc, Daniel
; APPLICANT: Goyette, Philippe
; APPLICANT: Campeau, Eric
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE, AND CANCER
; FILE REFERENCE: 50004/002005
; CURRENT APPLICATION NUMBER: US/10/607,712
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 08/980,326
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/031,964
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: 60/050,310
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-607-712-65

Query Match          0.6%; Score 25; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3788 CATTGAGGAAGAACATATCTGTGGC 3812
Db 25 CATTGAGGAAGAACATATCTGTGGC 1

RESULT 42
US-10-357-930-61072
; Sequence 61072, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61072
; LENGTH: 237
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-61072

Query Match          0.6%; Score 25; DB 8; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTTTCCTTTT 3887
```

```
Db 10 TTTTTCCTTTTTCCTTTTTCCTTTT 34

RESULT 43
US-10-424-599-9150/c
; Sequence 9150, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9150
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108272C.1
US-10-424-599-9150

Query Match          0.6%; Score 25; DB 7; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTTTCCTTTT 3887
Db 81 TTTTTCCTTTTTCCTTTTTCCTTTT 57

RESULT 44
US-10-424-599-13138
; Sequence 13138, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13138
; LENGTH: 262
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(262)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_111870C.1
US-10-424-599-13138

Query Match          0.6%; Score 25; DB 7; Length 262;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTTTCCTTTT 3887
Db 190 TTTTTCCTTTTTCCTTTTTCCTTTT 214

RESULT 45
```

US-10-425-115-158147  
; Sequence 158147, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 158147

; LENGTH: 280  
; TYPE: DNA  
; ORGANISM: Zea mays

; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_75807C.1  
US-10-425-115-158147

Query Match 0.6%; Score 25; DB 8; Length 280;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
|||||  
Db 6 TTTTGTGCTTTT 30

## RESULT 46

US-09-867-701-9532/c  
; Sequence 9532, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:

; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9532  
; LENGTH: 296  
; TYPE: DNA  
; ORGANISM: Homo sapien

US-09-867-701-9532

Query Match 0.6%; Score 25; DB 3; Length 296;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
|||||  
Db 288 TTTTGTGCTTTT 264

## RESULT 47

US-10-756-149-3961/c  
; Sequence 3961, Application US/10756149  
; Publication No. US20050181375A1  
; GENERAL INFORMATION:

; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10/756,149  
; CURRENT FILING DATE: 2004-01-12

; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3961  
; LENGTH: 343  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-756-149-3961

Query Match 0.6%; Score 25; DB 9; Length 343;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
|||||  
Db 329 TTTTGTGCTTTT 305

## RESULT 48

US-10-425-115-165833/c  
; Sequence 165833, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 165833

; LENGTH: 375  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_82824C.1  
US-10-425-115-165833

Query Match 0.6%; Score 25; DB 8; Length 375;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
|||||  
Db 162 TTTTGTGCTTTT 138

## RESULT 49

US-10-021-323-7632/c  
; Sequence 7632, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:

; APPLICANT: Deikman, Jill  
; APPLICANT: Peng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 7632

; LENGTH: 433  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3828-018-Q1-N6-F10  
US-10-021-323-7632

Query Match 0.6%; Score 25; DB 7; Length 433;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
Db 188 TTTTGTGCTTTT 164

RESULT 50  
US-10-425-115-111243/c  
; Sequence 111243, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 111243  
; LENGTH: 442  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_32945C.1  
US-10-425-115-111243

Query Match 0.6%; Score 25; DB 8; Length 442;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
Db 177 TTTTGTGCTTTT 153

Search completed: March 7, 2006, 03:38:15  
Job time : 2731 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 02:21:39 ; Search time 1025 Seconds  
(without alignments)  
8382.780 Million cell updates/sec

Title: US-10-607-712-1

Perfect score: 3919

Sequence: 1 ggtcaactgtggagagcagc.....ctcaagggaatacaacctag 3919

Scoring table: OIRGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 7218535 seqs, 1096242582 residues

Word size : 18

Total number of hits satisfying chosen parameters: 14824

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US08 NEW PUB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/US06 NEW PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US07 NEW PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/PCT NEW PUB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US09 NEW PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US09 NEW PUB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US10 NEW PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US10 NEW PUB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US11 NEW PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US11 NEW PUB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US11 NEW PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US11 NEW PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60 NEW PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	124	3.2	1275	6	US-09-925-065A-727226
C 2	80	2.0	616	6	US-09-925-065A-697175
C 3	39	1.0	611	6	US-09-925-065A-460077
C 4	31	0.8	588	6	US-09-925-065A-39453
5	26	0.7	559	6	US-09-925-065A-602656
6	26	0.7	595	6	US-09-925-065A-271845
7	26	0.7	1374	6	US-09-925-065A-702090
8	25	0.6	50	12	US-11-175-859-4373
C 9	25	0.6	542	6	US-09-925-065A-279223
C 10	25	0.6	545	6	US-09-925-065A-936651
C 11	25	0.6	630	6	US-09-925-065A-550047
12	25	0.6	1660	6	US-09-925-065A-34356
13	25	0.6	1660	6	US-09-925-065A-34357
14	25	0.6	1660	6	US-09-925-065A-34358
15	25	0.6	1660	6	US-09-925-065A-34359
16	25	0.6	1660	6	US-09-925-065A-34359
17	25	0.6	14911	12	US-11-112-908-63
C 18	25	0.6	156544	12	US-11-121-086-81
19	25	0.6	157230	12	US-11-112-908-64
20	25	0.6	170508	12	US-11-112-908-62

21	25	0.6	173115	12	US-11-112-908-65	Sequence 65, Appl
C 22	24	0.6	489	6	US-09-925-065A-412268	Sequence 412268,
C 23	24	0.6	506	6	US-09-925-065A-270554	Sequence 270554,
C 24	24	0.6	512	6	US-09-925-065A-620176	Sequence 620176,
25	24	0.6	548	6	US-09-925-065A-424567	Sequence 424567,
C 26	24	0.6	549	6	US-09-925-065A-889241	Sequence 889241,
C 27	24	0.6	585	6	US-09-925-065A-624935	Sequence 624935,
28	24	0.6	626	6	US-09-925-065A-530520	Sequence 530520,
C 29	24	0.6	241805	8	US-10-995-561-13215	Sequence 13215, A
C 30	24	0.6	387780	8	US-10-995-561-13259	Sequence 13259, A
31	23	0.6	25	8	US-10-310-914A-496602	Sequence 496602,
32	23	0.6	25	8	US-10-310-914A-859043	Sequence 859043,
33	23	0.6	201	12	US-11-124-367A-12473	Sequence 12473, A
34	23	0.6	201	12	US-11-124-367A-12474	Sequence 12474, A
C 35	23	0.6	376	6	US-09-925-065A-261182	Sequence 261182,
C 36	23	0.6	451	6	US-09-925-065A-351757	Sequence 351757,
C 37	23	0.6	481	8	US-10-689-742-18	Sequence 18, Appl
38	23	0.6	491	12	US-11-136-527-1395	Sequence 1395, Ap
C 39	23	0.6	491	12	US-11-136-527-5491	Sequence 5491, Ap
40	23	0.6	498	6	US-09-925-065A-98281	Sequence 98281, A
41	23	0.6	498	6	US-09-925-065A-98282	Sequence 98282, A
42	23	0.6	503	6	US-09-925-065A-925330	Sequence 925330,
C 43	23	0.6	522	6	US-09-925-065A-828324	Sequence 828324,
44	23	0.6	523	6	US-09-925-065A-819272	Sequence 819272,
45	23	0.6	556	6	US-09-925-065A-3410	Sequence 3410, Ap
C 46	23	0.6	574	6	US-09-925-065A-755600	Sequence 755600,
C 47	23	0.6	575	6	US-09-925-065A-278724	Sequence 278724,
48	23	0.6	576	6	US-09-925-065A-561677	Sequence 561677,
C 49	23	0.6	600	12	US-11-136-527-5990	Sequence 5990, Ap
C 50	23	0.6	601	6	US-09-925-065A-78484	Sequence 78484,
C 51	23	0.6	601	6	US-09-925-065A-478485	Sequence 478485,
52	23	0.6	611	6	US-09-925-065A-188252	Sequence 188252,
53	23	0.6	613	6	US-09-925-065A-751693	Sequence 751693,
54	23	0.6	634	6	US-09-925-065A-739010	Sequence 739010,
C 55	23	0.6	705	6	US-09-925-065A-922916	Sequence 922916,
56	23	0.6	955	6	US-09-925-065A-72630	Sequence 72630, A
C 57	23	0.6	1366	8	US-10-750-185-53615	Sequence 53615, A
C 58	23	0.6	1366	8	US-10-750-185-53615	Sequence 53615, A
C 59	23	0.6	1395	6	US-09-925-065A-78441	Sequence 78441, A
60	23	0.6	1629	8	US-10-750-185-45445	Sequence 45445, A
61	23	0.6	1629	8	US-10-750-185-45445	Sequence 45445, A
62	23	0.6	1736	8	US-10-750-185-64197	Sequence 64197, A
63	23	0.6	1736	8	US-10-750-185-64197	Sequence 64197, A
C 64	23	0.6	3077	8	US-10-750-185-64308	Sequence 64308, A
C 65	23	0.6	3077	8	US-10-750-185-64308	Sequence 64308, A
C 66	23	0.6	3188	8	US-10-750-185-55911	Sequence 55911, A
C 67	23	0.6	3188	8	US-10-750-185-55911	Sequence 55911, A
C 68	23	0.6	3434	12	US-11-136-527-1894	Sequence 1894, Ap
C 69	23	0.6	4913	8	US-10-750-185-59514	Sequence 59514, A
C 70	23	0.6	4913	8	US-10-750-185-59514	Sequence 59514, A
C 71	23	0.6	138821	12	US-11-121-086-80	Sequence 80, Appl
72	23	0.6	143947	12	US-11-193-561-37	Sequence 37, Appl
73	23	0.6	143947	12	US-11-193-771-37	Sequence 37, Appl
74	23	0.6	143947	12	US-11-193-789-37	Sequence 37, Appl
75	23	0.6	143947	12	US-11-193-806-37	Sequence 37, Appl
76	23	0.6	143947	12	US-11-193-857-37	Sequence 37, Appl
77	23	0.6	151169	12	US-11-121-086-38	Sequence 38, Appl
78	23	0.6	24	8	US-10-310-914A-202174	Sequence 202174,
79	22	0.6	24	8	US-10-310-914A-1050758	Sequence 1050758,
80	22	0.6	25	8	US-10-310-914A-1050759	Sequence 1050759,
81	22	0.6	95	8	US-10-995-561-26865	Sequence 26865, A
C 82	22	0.6	201	8	US-10-995-561-26985	Sequence 26985, A
C 83	22	0.6	201	8	US-10-995-561-26985	Sequence 26985, A
84	22	0.6	201	8	US-10-995-561-48316	Sequence 48316, A
C 85	22	0.6	201	12	US-11-124-367A-10359	Sequence 10359, A
C 86	22	0.6	201	12	US-11-124-367A-10360	Sequence 10360, A
C 87	22	0.6	374	12	US-11-198-847-4	Sequence 4, Appl1
C 88	22	0.6	424	6	US-09-925-065A-204738	Sequence 204738,
C 89	22	0.6	424	6	US-09-925-065A-204739	Sequence 204739,
C 90	22	0.6	481	6	US-09-925-065A-401865	Sequence 401865,
C 91	22	0.6	481	6	US-09-925-065A-401865	Sequence 401865,
C 92	22	0.6	494	6	US-09-925-065A-396325	Sequence 396325,
C 93	22	0.6	494	6	US-09-925-065A-396325	Sequence 396325,

94	22	0.6	505	6	US-09-925-065A-77965	Sequence 77965, A	167	21	0.5	26	8	US-10-310-914A-70474	Sequence 70474, A
c 95	22	0.6	514	6	US-09-925-065A-113707	Sequence 113707, A	168	21	0.5	26	8	US-10-310-914A-539759	Sequence 539759, A
c 96	22	0.6	519	6	US-09-925-065A-417525	Sequence 417525, A	169	21	0.5	26	8	US-10-310-914A-916377	Sequence 916377, A
c 97	22	0.6	519	6	US-09-925-065A-417527	Sequence 417527, A	c 170	21	0.5	31	8	US-10-310-914A-46964	Sequence 46964, A
c 98	22	0.6	520	6	US-09-925-065A-42413	Sequence 42413, A	c 171	21	0.5	31	8	US-10-310-914A-46965	Sequence 46965, A
c 99	22	0.6	524	6	US-09-925-065A-583132	Sequence 583132, A	c 172	21	0.5	201	8	US-10-995-561-15238	Sequence 15238, A
c 100	22	0.6	524	6	US-09-925-065A-650278	Sequence 650278, A	c 173	21	0.5	201	8	US-10-995-561-15240	Sequence 15240, A
c 101	22	0.6	524	6	US-09-925-065A-417523	Sequence 417523, A	c 174	21	0.5	201	8	US-10-995-561-23938	Sequence 23938, A
c 102	22	0.6	535	6	US-09-925-065A-417524	Sequence 417524, A	c 175	21	0.5	201	8	US-10-995-561-23940	Sequence 23940, A
c 103	22	0.6	535	6	US-09-925-065A-417524	Sequence 417524, A	c 176	21	0.5	201	8	US-10-995-561-24041	Sequence 24041, A
c 104	22	0.6	536	6	US-09-925-065A-924158	Sequence 924158, A	c 177	21	0.5	201	8	US-10-995-561-28047	Sequence 28047, A
c 105	22	0.6	538	6	US-09-925-065A-102798	Sequence 102798, A	c 178	21	0.5	201	8	US-10-995-561-32660	Sequence 32660, A
c 106	22	0.6	538	6	US-09-925-065A-102799	Sequence 102799, A	c 179	21	0.5	201	8	US-10-995-561-40940	Sequence 40940, A
c 107	22	0.6	540	6	US-09-925-065A-11434	Sequence 11434, A	c 180	21	0.5	201	8	US-10-995-561-42325	Sequence 42325, A
c 108	22	0.6	540	6	US-09-925-065A-105964	Sequence 105964, A	c 181	21	0.5	201	12	US-11-124-367A-5382	Sequence 5382, A
c 109	22	0.6	550	6	US-09-925-065A-262968	Sequence 262968, A	c 182	21	0.5	201	12	US-11-124-367A-5704	Sequence 5704, A
c 110	22	0.6	563	6	US-09-925-065A-898853	Sequence 898853, A	c 183	21	0.5	201	12	US-11-124-367A-28717	Sequence 28717, A
c 111	22	0.6	569	6	US-09-925-065A-112247	Sequence 112247, A	c 184	21	0.5	201	12	US-11-124-367A-28742	Sequence 28742, A
c 112	22	0.6	569	6	US-09-925-065A-112248	Sequence 112248, A	c 185	21	0.5	285	6	US-09-925-065A-130263	Sequence 130263, A
c 113	22	0.6	570	6	US-09-925-065A-314680	Sequence 314680, A	c 186	21	0.5	345	6	US-09-925-065A-732590	Sequence 732590, A
c 114	22	0.6	578	6	US-09-925-065A-314681	Sequence 314681, A	c 187	21	0.5	353	6	US-09-925-065A-149005	Sequence 149005, A
c 115	22	0.6	589	6	US-09-925-065A-297980	Sequence 297980, A	c 188	21	0.5	372	6	US-09-925-065A-554264	Sequence 554264, A
c 116	22	0.6	598	6	US-09-925-065A-878937	Sequence 878937, A	c 189	21	0.5	398	6	US-09-925-065A-177621	Sequence 177621, A
c 117	22	0.6	598	6	US-09-925-065A-539793	Sequence 539793, A	c 190	21	0.5	398	6	US-09-925-065A-177622	Sequence 177622, A
c 118	22	0.6	600	6	US-09-925-065A-902706	Sequence 902706, A	c 191	21	0.5	415	6	US-09-925-065A-518970	Sequence 518970, A
c 119	22	0.6	600	8	US-10-750-185A-4706	Sequence 4706, A	c 192	21	0.5	415	6	US-09-925-065A-518971	Sequence 518971, A
c 120	22	0.6	600	12	US-11-136-527-4665	Sequence 4665, A	c 193	21	0.5	424	6	US-09-925-065A-219497	Sequence 219497, A
c 121	22	0.6	601	6	US-09-925-065A-575787	Sequence 575787, A	c 194	21	0.5	427	6	US-09-925-065A-154328	Sequence 154328, A
c 122	22	0.6	605	6	US-09-925-065A-898473	Sequence 898473, A	c 195	21	0.5	445	6	US-09-925-065A-349522	Sequence 349522, A
c 123	22	0.6	608	6	US-09-925-065A-637147	Sequence 637147, A	c 196	21	0.5	456	6	US-09-925-065A-513522	Sequence 513522, A
c 124	22	0.6	609	6	US-09-925-065A-809765	Sequence 809765, A	c 197	21	0.5	456	6	US-09-925-065A-561045	Sequence 561045, A
c 125	22	0.6	635	6	US-09-925-065A-731865	Sequence 731865, A	c 198	21	0.5	456	6	US-09-925-065A-561047	Sequence 561047, A
c 126	22	0.6	734	6	US-09-925-065A-80511	Sequence 80511, A	c 199	21	0.5	456	6	US-09-925-065A-561048	Sequence 561048, A
c 127	22	0.6	734	6	US-09-925-065A-80512	Sequence 80512, A	c 200	21	0.5	471	6	US-09-925-065A-496565	Sequence 496565, A
c 128	22	0.6	817	12	US-11-136-527-569	Sequence 569, A	c 201	21	0.5	471	6	US-09-925-065A-496566	Sequence 496566, A
c 129	22	0.6	952	6	US-09-925-065A-101979	Sequence 101979, A	c 202	21	0.5	485	12	US-11-198-847-220	Sequence 220, A
c 130	22	0.6	980	6	US-09-925-065A-696097	Sequence 696097, A	c 203	21	0.5	489	6	US-09-925-065A-289154	Sequence 289154, A
c 131	22	0.6	982	6	US-09-925-065A-725590	Sequence 725590, A	c 204	21	0.5	496	6	US-09-925-065A-753404	Sequence 753404, A
c 132	22	0.6	1170	6	US-09-925-065A-297981	Sequence 297981, A	c 205	21	0.5	496	6	US-09-925-065A-753405	Sequence 753405, A
c 133	22	0.6	1225	6	US-09-925-065A-703263	Sequence 703263, A	c 206	21	0.5	497	6	US-09-925-065A-260244	Sequence 260244, A
c 134	22	0.6	1573	8	US-10-689-742-187	Sequence 187, A	c 207	21	0.5	500	6	US-09-925-065A-522553	Sequence 522553, A
c 135	22	0.6	1599	8	US-10-750-185-28515	Sequence 28515, A	c 208	21	0.5	503	6	US-09-925-065A-387021	Sequence 387021, A
c 136	22	0.6	1599	8	US-10-750-623-28515	Sequence 28515, A	c 209	21	0.5	503	6	US-09-925-065A-387022	Sequence 387022, A
c 137	22	0.6	1736	8	US-10-750-185-35826	Sequence 35826, A	c 210	21	0.5	505	6	US-09-925-065A-751623	Sequence 751623, A
c 138	22	0.6	1736	8	US-10-750-623-35826	Sequence 35826, A	c 211	21	0.5	510	6	US-09-925-065A-405485	Sequence 405485, A
c 139	22	0.6	1766	6	US-09-925-065A-30791	Sequence 30791, A	c 212	21	0.5	510	6	US-09-925-065A-405486	Sequence 405486, A
c 140	22	0.6	2315	6	US-09-925-065A-706810	Sequence 706810, A	c 213	21	0.5	511	6	US-09-925-065A-74198	Sequence 74198, A
c 141	22	0.6	2315	6	US-09-925-065A-706811	Sequence 706811, A	c 214	21	0.5	511	6	US-09-925-065A-74199	Sequence 74199, A
c 142	22	0.6	2628	6	US-09-925-065A-676638	Sequence 676638, A	c 215	21	0.5	511	6	US-09-925-065A-74200	Sequence 74200, A
c 143	22	0.6	2828	12	US-11-004-762-22	Sequence 22, A	c 216	21	0.5	512	6	US-09-925-065A-580299	Sequence 580299, A
c 144	22	0.6	3002	6	US-09-925-065A-674422	Sequence 674422, A	c 217	21	0.5	512	6	US-09-925-065A-762931	Sequence 762931, A
c 145	22	0.6	3002	6	US-09-925-065A-674423	Sequence 674423, A	c 218	21	0.5	513	6	US-09-925-065A-358459	Sequence 358459, A
c 146	22	0.6	4216	8	US-10-750-185-31675	Sequence 31675, A	c 219	21	0.5	513	6	US-09-925-065A-358460	Sequence 358460, A
c 147	22	0.6	4216	8	US-10-750-623-31675	Sequence 31675, A	c 220	21	0.5	513	6	US-09-925-065A-358461	Sequence 358461, A
c 148	22	0.6	4608	8	US-10-750-185-27329	Sequence 27329, A	c 221	21	0.5	513	6	US-09-925-065A-358462	Sequence 358462, A
c 149	22	0.6	4608	8	US-10-750-623-27329	Sequence 27329, A	c 222	21	0.5	513	6	US-09-925-065A-763143	Sequence 763143, A
c 150	22	0.6	6450	7	US-10-826-585-21	Sequence 21, A	c 223	21	0.5	514	6	US-09-925-065A-484457	Sequence 484457, A
c 151	22	0.6	6450	8	US-10-955-054A-30	Sequence 30, A	c 224	21	0.5	521	6	US-09-925-065A-422732	Sequence 422732, A
c 152	22	0.6	22475	7	US-10-330-773-975	Sequence 975, A	c 225	21	0.5	524	6	US-09-925-065A-118032	Sequence 118032, A
c 153	22	0.6	79528	8	US-10-276-233A-6	Sequence 6, A	c 226	21	0.5	526	6	US-09-925-065A-320969	Sequence 320969, A
c 154	22	0.6	81210	8	US-10-995-561-13295	Sequence 13295, A	c 227	21	0.5	526	6	US-09-925-065A-320970	Sequence 320970, A
c 155	22	0.6	87672	8	US-10-995-561-13237	Sequence 13237, A	c 228	21	0.5	530	6	US-09-925-065A-318951	Sequence 318951, A
c 156	22	0.6	171732	12	US-11-121-086-98	Sequence 98, A	c 229	21	0.5	530	6	US-09-925-065A-318953	Sequence 318953, A
c 157	22	0.6	178024	7	US-10-330-773-698	Sequence 698, A	c 230	21	0.5	530	6	US-09-925-065A-318954	Sequence 318954, A
c 158	22	0.6	318488	9	US-11-114-798-58	Sequence 58, A	c 231	21	0.5	530	6	US-09-925-065A-557580	Sequence 557580, A
c 159	21	0.5	21	8	US-10-310-914A-263153	Sequence 263153, A	c 232	21	0.5	532	6	US-09-925-065A-610344	Sequence 610344, A
c 160	21	0.5	22	8	US-10-310-914A-70473	Sequence 70473, A	c 233	21	0.5	532	6	US-09-925-065A-610345	Sequence 610345, A
c 161	21	0.5	22	8	US-10-310-914A-539758	Sequence 539758, A	c 234	21	0.5	532	6	US-09-925-065A-610346	Sequence 610346, A
c 162	21	0.5	22	8	US-10-310-914A-856782	Sequence 856782, A	c 235	21	0.5	532	6	US-09-925-065A-610347	Sequence 610347, A
c 163	21	0.5	22	8	US-10-310-914A-862411	Sequence 862411, A	c 236	21	0.5	533	6	US-09-925-065A-526099	Sequence 526099, A
c 164	21	0.5	22	8	US-10-310-914A-882783	Sequence 882783, A	c 237	21	0.5	534	6	US-09-925-065A-10121	Sequence 10121, A
c 165	21	0.5	22	8	US-10-310-914A-916376	Sequence 916376, A	c 238	21	0.5	534	6	US-09-925-065A-333898	Sequence 333898, A
c 166	21	0.5	23	8	US-10-310-914A-882784	Sequence 882784, A	c 239	21	0.5	537	6	US-09-925-065A-230349	Sequence 230349, A

240	21	0.5	537	6	US-09-925-065A-489872	Sequence 489872,	c 313	21	0.5	627	6	US-09-925-065A-878743	Sequence 878743,
241	21	0.5	537	6	US-09-925-065A-489873	Sequence 489873,	c 314	21	0.5	627	6	US-09-925-065A-878744	Sequence 878744,
242	21	0.5	537	6	US-09-925-065A-489874	Sequence 489874,	c 315	21	0.5	627	6	US-09-925-065A-878745	Sequence 878745,
c 243	21	0.5	538	6	US-09-925-065A-340190	Sequence 340190,	c 316	21	0.5	628	6	US-09-925-065A-559335	Sequence 559335,
244	21	0.5	546	6	US-09-925-065A-711178	Sequence 711178, A	c 317	21	0.5	633	6	US-09-925-065A-875899	Sequence 875899,
245	21	0.5	546	6	US-09-925-065A-711179	Sequence 711179, A	c 318	21	0.5	636	6	US-09-925-065A-731134	Sequence 731134,
c 246	21	0.5	547	6	US-09-925-065A-589347	Sequence 589347,	c 319	21	0.5	636	6	US-09-925-065A-731135	Sequence 731135,
247	21	0.5	550	6	US-09-925-065A-319926	Sequence 319926,	c 320	21	0.5	637	6	US-09-925-065A-932840	Sequence 932840,
c 248	21	0.5	550	6	US-09-925-065A-319927	Sequence 319927,	c 321	21	0.5	638	6	US-09-925-065A-548422	Sequence 548422,
c 249	21	0.5	556	6	US-09-925-065A-591297	Sequence 591297,	c 322	21	0.5	639	6	US-09-925-065A-794334	Sequence 794334,
250	21	0.5	557	6	US-09-925-065A-123387	Sequence 123387,	c 323	21	0.5	639	6	US-09-925-065A-822586	Sequence 822586,
c 251	21	0.5	558	6	US-09-925-065A-293387	Sequence 293387,	c 324	21	0.5	640	6	US-09-925-065A-688042	Sequence 688042,
c 252	21	0.5	558	6	US-09-925-065A-454776	Sequence 454776,	c 325	21	0.5	640	6	US-09-925-065A-688043	Sequence 688043,
c 253	21	0.5	558	6	US-09-925-065A-454777	Sequence 454777, A	c 326	21	0.5	640	6	US-09-925-065A-862829	Sequence 862829,
c 254	21	0.5	560	6	US-09-925-065A-17376	Sequence 17376, A	c 327	21	0.5	640	6	US-09-925-065A-862830	Sequence 862830,
255	21	0.5	562	6	US-09-925-065A-158888	Sequence 158888,	c 328	21	0.5	640	6	US-09-925-065A-897747	Sequence 897747,
256	21	0.5	575	6	US-09-925-065A-938868	Sequence 938868,	c 329	21	0.5	650	6	US-09-925-065A-701394	Sequence 701394,
257	21	0.5	575	6	US-09-925-065A-938869	Sequence 938869,	c 330	21	0.5	653	6	US-09-925-065A-748516	Sequence 748516,
c 258	21	0.5	577	6	US-09-925-065A-258961	Sequence 258961,	c 331	21	0.5	653	6	US-09-925-065A-748517	Sequence 748517,
c 259	21	0.5	577	6	US-09-925-065A-258962	Sequence 258962,	c 332	21	0.5	653	6	US-09-925-065A-821566	Sequence 821566,
c 260	21	0.5	579	6	US-09-925-065A-954549	Sequence 954549,	c 333	21	0.5	657	6	US-09-925-065A-41772	Sequence 41772, A
c 261	21	0.5	580	6	US-09-925-065A-296463	Sequence 296463,	c 334	21	0.5	657	6	US-09-925-065A-41773	Sequence 41773, A
262	21	0.5	582	6	US-09-925-065A-899637	Sequence 899637,	c 335	21	0.5	668	6	US-09-925-065A-691735	Sequence 691735,
263	21	0.5	582	6	US-09-925-065A-899638	Sequence 899638,	c 336	21	0.5	672	6	US-09-925-065A-682472	Sequence 682472,
c 264	21	0.5	583	6	US-09-925-065A-40758	Sequence 40758, A	c 337	21	0.5	673	6	US-09-925-065A-680053	Sequence 680053,
c 265	21	0.5	583	6	US-09-925-065A-508657	Sequence 508657,	c 338	21	0.5	700	6	US-09-925-065A-540081	Sequence 540081,
c 266	21	0.5	585	6	US-09-925-065A-172098	Sequence 172098,	c 339	21	0.5	710	6	US-09-925-065A-65701	Sequence 65701, A
c 267	21	0.5	585	6	US-09-925-065A-172099	Sequence 172099,	c 340	21	0.5	726	6	US-09-925-065A-63027	Sequence 63027, A
c 268	21	0.5	585	6	US-09-925-065A-172100	Sequence 172100,	c 341	21	0.5	732	6	US-09-925-065A-57310	Sequence 57310, A
c 269	21	0.5	585	6	US-09-925-065A-294385	Sequence 294385,	c 342	21	0.5	762	6	US-09-925-065A-59562	Sequence 59562, A
c 270	21	0.5	585	6	US-09-925-065A-294386	Sequence 294386,	c 343	21	0.5	762	6	US-09-925-065A-59563	Sequence 59563, A
c 271	21	0.5	585	6	US-09-925-065A-774157	Sequence 774157,	c 344	21	0.5	801	6	US-09-925-065A-678775	Sequence 678775,
c 272	21	0.5	587	6	US-09-925-065A-774158	Sequence 774158,	c 345	21	0.5	895	8	US-10-750-185-61125	Sequence 61125, A
c 273	21	0.5	587	6	US-09-925-065A-670881	Sequence 670881,	c 346	21	0.5	895	8	US-10-750-623-61125	Sequence 61125, A
c 274	21	0.5	587	6	US-09-925-065A-670882	Sequence 670882,	c 347	21	0.5	907	6	US-09-925-065A-68284	Sequence 68284, A
c 275	21	0.5	587	6	US-09-925-065A-670883	Sequence 670883,	c 348	21	0.5	907	6	US-09-925-065A-68285	Sequence 68285, A
c 276	21	0.5	589	6	US-09-925-065A-588995	Sequence 588995,	c 349	21	0.5	907	6	US-09-925-065A-31170	Sequence 31170, A
c 277	21	0.5	589	6	US-09-925-065A-588996	Sequence 588996,	c 350	21	0.5	1002	6	US-09-925-065A-3659	Sequence 3659, Ap
c 278	21	0.5	589	6	US-09-925-065A-588997	Sequence 588997,	c 351	21	0.5	1027	6	US-09-925-065A-3660	Sequence 3660, Ap
c 279	21	0.5	592	6	US-09-925-065A-895820	Sequence 895820,	c 352	21	0.5	1027	6	US-09-925-065A-3661	Sequence 3661, Ap
c 280	21	0.5	593	6	US-09-925-065A-895821	Sequence 895821,	c 353	21	0.5	1027	6	US-09-925-065A-3662	Sequence 3662, Ap
c 281	21	0.5	595	6	US-09-925-065A-886355	Sequence 886355,	c 354	21	0.5	1027	6	US-09-925-065A-3662	Sequence 3662, Ap
c 282	21	0.5	595	6	US-09-925-065A-911361	Sequence 911361,	c 355	21	0.5	1070	6	US-09-925-065A-686010	Sequence 686010,
c 283	21	0.5	597	6	US-09-925-065A-762930	Sequence 762930,	c 356	21	0.5	1070	6	US-09-925-065A-686011	Sequence 686011,
c 284	21	0.5	598	6	US-09-925-065A-816088	Sequence 816088,	c 357	21	0.5	1087	6	US-09-925-065A-709482	Sequence 709482,
c 285	21	0.5	598	6	US-09-925-065A-399183	Sequence 399183,	c 358	21	0.5	1087	6	US-09-925-065A-709483	Sequence 709483,
c 286	21	0.5	599	6	US-09-925-065A-399184	Sequence 399184,	c 359	21	0.5	1092	6	US-09-925-065A-689194	Sequence 689194,
c 287	21	0.5	600	6	US-09-925-065A-690628	Sequence 690628,	c 360	21	0.5	1130	6	US-09-925-065A-719120	Sequence 719120,
c 288	21	0.5	603	6	US-09-925-065A-319246	Sequence 319246,	c 361	21	0.5	1130	6	US-09-925-065A-719121	Sequence 719121,
c 289	21	0.5	610	6	US-09-925-065A-580587	Sequence 580587,	c 362	21	0.5	1131	6	US-09-925-065A-718023	Sequence 718023,
c 290	21	0.5	611	6	US-09-925-065A-885141	Sequence 885141,	c 363	21	0.5	1131	6	US-09-925-065A-718024	Sequence 718024,
c 291	21	0.5	611	6	US-09-925-065A-910826	Sequence 910826,	c 364	21	0.5	1139	6	US-09-925-065A-28729	Sequence 28729, A
c 292	21	0.5	612	6	US-09-925-065A-161287	Sequence 161287,	c 365	21	0.5	1139	6	US-09-925-065A-28730	Sequence 28730, A
c 293	21	0.5	612	6	US-09-925-065A-161288	Sequence 161288,	c 366	21	0.5	1139	6	US-09-925-065A-28731	Sequence 28731, A
c 294	21	0.5	612	6	US-09-925-065A-247900	Sequence 247900,	c 367	21	0.5	1139	6	US-09-925-065A-28732	Sequence 28732, A
c 295	21	0.5	612	6	US-09-925-065A-343174	Sequence 343174,	c 368	21	0.5	1167	6	US-09-925-065A-670666	Sequence 670666,
c 296	21	0.5	612	6	US-09-925-065A-343175	Sequence 343175,	c 369	21	0.5	1190	6	US-09-925-065A-704144	Sequence 704144,
c 297	21	0.5	612	6	US-09-925-065A-343176	Sequence 343176,	c 370	21	0.5	1284	6	US-09-925-065A-681170	Sequence 681170,
c 298	21	0.5	613	6	US-09-925-065A-928668	Sequence 928668,	c 371	21	0.5	1284	6	US-09-925-065A-681171	Sequence 681171,
c 299	21	0.5	615	6	US-09-925-065A-770572	Sequence 770572,	c 372	21	0.5	1284	6	US-09-925-065A-681172	Sequence 681172,
c 300	21	0.5	619	6	US-09-925-065A-47436	Sequence 47436, A	c 373	21	0.5	1331	6	US-09-925-065A-47638	Sequence 47638, A
c 301	21	0.5	619	6	US-09-925-065A-556506	Sequence 556506,	c 374	21	0.5	1332	6	US-09-925-065A-69581	Sequence 69581, A
c 302	21	0.5	620	6	US-09-925-065A-866942	Sequence 866942,	c 375	21	0.5	1382	6	US-09-925-065A-69582	Sequence 69582, A
c 303	21	0.5	620	6	US-09-925-065A-866943	Sequence 866943,	c 376	21	0.5	1382	6	US-09-925-065A-69583	Sequence 69583, A
c 304	21	0.5	620	6	US-09-925-065A-899725	Sequence 899725,	c 377	21	0.5	1382	6	US-09-925-065A-69584	Sequence 69584, A
c 305	21	0.5	622	6	US-09-925-065A-501311	Sequence 501311,	c 378	21	0.5	1382	6	US-09-925-065A-550969	Sequence 550969,
c 306	21	0.5	623	6	US-09-925-065A-501312	Sequence 501312,	c 379	21	0.5	1393	8	US-10-750-185-54425	Sequence 54425, A
c 307	21	0.5	623	6	US-09-925-065A-264948	Sequence 264948,	c 380	21	0.5	1393	8	US-10-750-623-54425	Sequence 54425, A
c 308	21	0.5	625	6	US-09-925-065A-15616	Sequence 15616, A	c 381	21	0.5	1398	12	US-11-149-403-3	Sequence 3, Appl1
c 309	21	0.5	625	6	US-09-925-065A-158695	Sequence 158695,	c 382	21	0.5	1496	8	US-10-750-185-55271	Sequence 55271, A
c 310	21	0.5	625	6	US-09-925-065A-158696	Sequence 158696,	c 383	21	0.5	1496	8	US-10-750-623-55271	Sequence 55271, A
c 311	21	0.5	626	6	US-09-925-065A-938813	Sequence 938813,	c 384	21	0.5	1512	6	US-09-925-065A-704680	Sequence 704680,
c 312	21	0.5	626	6	US-09-925-065A-938814	Sequence 938814,	c 385	21	0.5	1512	6	US-09-925-065A-704681	Sequence 704681,

C 386	21	0.5	1673	6	US-09-925-065A-73000	Sequence 73000, A	C 459	21	0.5	170508	12	US-11-112-908-62	Sequence 62, Appl
C 387	21	0.5	1694	6	US-09-925-065A-699512	Sequence 699512, A	C 460	21	0.5	171162	12	US-11-112-908-38	Sequence 38, Appl
C 388	21	0.5	1694	6	US-09-925-065A-699513	Sequence 699513, A	C 461	21	0.5	171162	12	US-11-112-908-65	Sequence 65, Appl
C 389	21	0.5	1760	6	US-09-925-065A-714280	Sequence 714280, A	C 462	21	0.5	173120	9	US-11-114-798-55	Sequence 55, Appl
C 390	21	0.5	1761	8	US-10-750-185-32845	Sequence 32845, A	C 463	21	0.5	176503	12	US-11-121-086-53	Sequence 53, Appl
C 391	21	0.5	1761	8	US-10-750-623-32845	Sequence 32845, A	C 464	21	0.5	178024	7	US-10-330-773-698	Sequence 698, App
C 392	21	0.5	1774	8	US-10-750-185-62002	Sequence 62002, A	C 465	21	0.5	182303	12	US-11-121-086-45	Sequence 45, Appl
C 393	21	0.5	1774	8	US-10-750-623-62002	Sequence 62002, A	C 466	21	0.5	196200	12	US-11-121-086-9	Sequence 9, Appl
C 394	21	0.5	1778	8	US-10-980-388-54	Sequence 54, Appl	C 467	21	0.5	197241	9	US-11-114-798-47	Sequence 47, Appl
C 395	21	0.5	1798	8	US-09-925-065A-685610	Sequence 685610, A	C 468	21	0.5	218821	12	US-11-121-086-31	Sequence 31, Appl
C 396	21	0.5	1824	8	US-10-750-185-55137	Sequence 55137, A	C 469	21	0.5	222094	8	US-10-995-561-13244	Sequence 13244, A
C 397	21	0.5	1824	8	US-10-750-623-55137	Sequence 55137, A	C 470	21	0.5	235033	12	US-11-157-389-1	Sequence 1, Appl
C 398	21	0.5	1831	8	US-10-750-185-34142	Sequence 34142, A	C 471	21	0.5	237326	12	US-11-157-389-2	Sequence 2, Appl
C 399	21	0.5	1831	8	US-10-750-623-34142	Sequence 34142, A	C 472	21	0.5	246960	12	US-11-121-086-8	Sequence 8, Appl
C 400	21	0.5	1962	8	US-10-750-185-31991	Sequence 31991, A	C 473	21	0.5	257645	7	US-10-330-773-266	Sequence 266, App
C 401	21	0.5	1962	8	US-10-750-623-31991	Sequence 31991, A	C 474	21	0.5	355211	7	US-10-330-773-242	Sequence 242, App
C 402	21	0.5	2195	6	US-09-925-065A-13945	Sequence 13945, A	C 475	21	0.5	440684	7	US-10-330-773-114	Sequence 114, App
C 403	21	0.5	2195	6	US-09-925-065A-13946	Sequence 13946, A	C 476	21	0.5	440684	7	US-10-330-773-114	Sequence 114, App
C 404	21	0.5	2260	6	US-09-925-065A-690385	Sequence 690385, A	C 477	21	0.5	1125000	8	US-10-995-561-13286	Sequence 13286, A
C 405	21	0.5	2260	6	US-09-925-065A-690386	Sequence 690386, A	C 478	21	0.5	1125000	8	US-10-995-561-13286	Sequence 13286, A
C 406	21	0.5	2277	6	US-09-925-065A-685269	Sequence 685269, A	C 479	20	0.5	20	8	US-10-310-914A-46760	Sequence 46760, A
C 407	21	0.5	2277	6	US-09-925-065A-685270	Sequence 685270, A	C 480	20	0.5	20	8	US-10-310-914A-979695	Sequence 979695, A
C 408	21	0.5	2277	6	US-09-925-065A-685271	Sequence 685271, A	C 481	20	0.5	21	8	US-10-310-914A-84873	Sequence 84873, A
C 409	21	0.5	2369	6	US-09-925-065A-707978	Sequence 707978, A	C 482	20	0.5	21	8	US-10-310-914A-191734	Sequence 191734, A
C 410	21	0.5	2369	6	US-09-925-065A-707979	Sequence 707979, A	C 483	20	0.5	21	8	US-10-310-914A-191734	Sequence 191734, A
C 411	21	0.5	2369	6	US-09-925-065A-707980	Sequence 707980, A	C 484	20	0.5	21	8	US-10-310-914A-224007	Sequence 224007, A
C 412	21	0.5	2369	6	US-09-925-065A-707981	Sequence 707981, A	C 485	20	0.5	21	8	US-10-310-914A-432268	Sequence 432268, A
C 413	21	0.5	2369	6	US-09-925-065A-707982	Sequence 707982, A	C 486	20	0.5	21	8	US-10-310-914A-454705	Sequence 454705, A
C 414	21	0.5	2369	6	US-09-925-065A-707983	Sequence 707983, A	C 487	20	0.5	21	8	US-10-310-914A-466979	Sequence 466979, A
C 415	21	0.5	2369	6	US-09-925-065A-707984	Sequence 707984, A	C 488	20	0.5	21	8	US-10-310-914A-661429	Sequence 661429, A
C 416	21	0.5	2465	9	US-11-072-512-1278	Sequence 1278, Ap	C 489	20	0.5	21	8	US-10-310-914A-670173	Sequence 670173, A
C 417	21	0.5	2498	6	US-09-925-065A-685840	Sequence 685840, A	C 490	20	0.5	21	8	US-10-310-914A-698735	Sequence 698735, A
C 418	21	0.5	2498	6	US-09-925-065A-685841	Sequence 685841, A	C 491	20	0.5	21	8	US-10-310-914A-698736	Sequence 698736, A
C 419	21	0.5	2723	8	US-10-750-185-54580	Sequence 54580, A	C 492	20	0.5	21	8	US-10-310-914A-877571	Sequence 877571, A
C 420	21	0.5	2731	9	US-11-072-512-284	Sequence 284, App	C 493	20	0.5	21	8	US-10-310-914A-1335405	Sequence 1335405, A
C 421	21	0.5	2731	9	US-09-925-065A-677497	Sequence 677497, A	C 494	20	0.5	22	8	US-10-310-914A-179441	Sequence 179441, A
C 422	21	0.5	2765	6	US-09-925-065A-723375	Sequence 723375, A	C 495	20	0.5	22	8	US-10-310-914A-263297	Sequence 263297, A
C 423	21	0.5	2941	6	US-09-925-065A-717788	Sequence 717788, A	C 496	20	0.5	22	8	US-10-310-914A-670172	Sequence 670172, A
C 424	21	0.5	3227	9	US-11-245-147-112	Sequence 112, App	C 497	20	0.5	22	8	US-10-310-914A-906435	Sequence 906435, A
C 425	21	0.5	3576	8	US-10-750-185-33942	Sequence 33942, A	C 498	20	0.5	22	8	US-10-310-914A-979696	Sequence 979696, A
C 426	21	0.5	3576	8	US-10-750-623-33942	Sequence 33942, A	C 499	20	0.5	23	8	US-10-310-914A-84871	Sequence 84871, A
C 427	21	0.5	3711	6	US-09-925-065A-717787	Sequence 717787, A	C 500	20	0.5	23	8	US-10-310-914A-205410	Sequence 205410, A
C 428	21	0.5	3711	6	US-09-925-065A-717788	Sequence 717788, A	C 501	20	0.5	23	8	US-10-310-914A-877572	Sequence 877572, A
C 429	21	0.5	3711	6	US-09-925-065A-717789	Sequence 717789, A	C 502	20	0.5	23	8	US-10-310-914A-179444	Sequence 179444, A
C 430	21	0.5	3711	6	US-09-925-065A-717790	Sequence 717790, A	C 503	20	0.5	24	8	US-10-310-914A-191735	Sequence 191735, A
C 431	21	0.5	4105	8	US-10-750-185-37015	Sequence 37015, A	C 504	20	0.5	24	8	US-10-310-914A-454706	Sequence 454706, A
C 432	21	0.5	4105	8	US-10-750-623-37015	Sequence 37015, A	C 505	20	0.5	24	8	US-10-310-914A-454707	Sequence 454707, A
C 433	21	0.5	4295	8	US-10-750-185-55549	Sequence 55549, A	C 506	20	0.5	24	8	US-10-310-914A-466980	Sequence 466980, A
C 434	21	0.5	4295	8	US-10-750-623-55549	Sequence 55549, A	C 507	20	0.5	24	8	US-10-310-914A-698737	Sequence 698737, A
C 435	21	0.5	4619	8	US-10-750-185-57460	Sequence 57460, A	C 508	20	0.5	24	8	US-10-310-914A-877569	Sequence 877569, A
C 436	21	0.5	4619	8	US-10-750-623-57460	Sequence 57460, A	C 509	20	0.5	24	8	US-10-310-914A-1090902	Sequence 1090902, A
C 437	21	0.5	5585	8	US-10-750-185-47879	Sequence 47879, A	C 510	20	0.5	24	8	US-10-310-914A-1335406	Sequence 1335406, A
C 438	21	0.5	5585	8	US-10-750-623-47879	Sequence 47879, A	C 511	20	0.5	25	8	US-10-310-914A-432269	Sequence 432269, A
C 439	21	0.5	10467	8	US-10-240-708-1	Sequence 1, Appl	C 512	20	0.5	26	8	US-10-310-914A-661430	Sequence 661430, A
C 440	21	0.5	20317	7	US-10-995-561-13460	Sequence 13460, A	C 513	20	0.5	201	8	US-10-995-561-13577	Sequence 13577, A
C 441	21	0.5	26354	7	US-10-330-773-444	Sequence 444, App	C 514	20	0.5	201	8	US-10-995-561-13832	Sequence 13832, A
C 442	21	0.5	27902	8	US-10-995-561-13462	Sequence 13462, A	C 515	20	0.5	201	8	US-10-995-561-22749	Sequence 22749, A
C 443	21	0.5	69531	8	US-10-995-561-13254	Sequence 13254, A	C 516	20	0.5	201	8	US-10-995-561-26422	Sequence 26422, A
C 444	21	0.5	69648	7	US-10-330-773-911	Sequence 911, App	C 517	20	0.5	201	8	US-10-995-561-26659	Sequence 26659, A
C 445	21	0.5	77246	12	US-11-124-368A-2907	Sequence 2907, Ap	C 518	20	0.5	201	8	US-10-995-561-32124	Sequence 32124, A
C 446	21	0.5	98638	7	US-10-330-773-896	Sequence 896, App	C 519	20	0.5	201	8	US-10-995-561-32125	Sequence 32125, A
C 447	21	0.5	100000	12	US-11-124-368A-2890	Sequence 2890, Ap	C 520	20	0.5	201	8	US-10-995-561-36928	Sequence 36928, A
C 448	21	0.5	101169	7	US-10-330-773-561	Sequence 561, App	C 521	20	0.5	201	8	US-10-995-561-39335	Sequence 39335, A
C 449	21	0.5	124285	7	US-10-893-483-105	Sequence 105, App	C 522	20	0.5	201	8	US-10-995-561-39704	Sequence 39704, A
C 450	21	0.5	127722	7	US-10-330-773-278	Sequence 278, App	C 523	20	0.5	201	8	US-10-995-561-39818	Sequence 39818, A
C 451	21	0.5	151052	7	US-10-330-773-503	Sequence 503, App	C 524	20	0.5	201	8	US-10-995-561-40197	Sequence 40197, A
C 452	21	0.5	151870	7	US-10-995-561-13199	Sequence 13199, A	C 525	20	0.5	201	8	US-10-995-561-40250	Sequence 40250, A
C 453	21	0.5	152096	7	US-10-330-773-942	Sequence 942, App	C 526	20	0.5	201	8	US-10-995-561-42961	Sequence 42961, A
C 454	21	0.5	157230	12	US-11-112-908-64	Sequence 64, Appl	C 527	20	0.5	201	8	US-10-995-561-44270	Sequence 44270, A
C 455	21	0.5	166639	12	US-11-121-086-52	Sequence 52, Appl	C 528	20	0.5	201	8	US-10-995-561-44527	Sequence 44527, A
C 456	21	0.5	168516	12	US-11-121-086-3	Sequence 3, Appl	C 529	20	0.5	201	8	US-10-995-561-46158	Sequence 46158, A
C 457	21	0.5	168656	12	US-11-112-908-59	Sequence 59, Appl	C 530	20	0.5	201	8	US-10-995-561-49557	Sequence 49557, A
C 458	21	0.5	170285	12	US-11-112-908-58	Sequence 58, Appl	C 531	20	0.5	201	8	US-10-995-561-52271	Sequence 52271, A

C 532	20	0.5	201	8	US-10-995-561-52622	Sequence 52622, A	C 605	20	0.5	333	6	US-09-925-065A-576379	Sequence 576379,
C 533	20	0.5	201	8	US-10-995-561-53793	Sequence 53793, A	C 606	20	0.5	336	6	US-09-925-065A-370032	Sequence 370032,
C 534	20	0.5	201	8	US-10-995-561-53851	Sequence 53851, A	C 607	20	0.5	337	6	US-09-925-065A-176025	Sequence 176025,
C 535	20	0.5	201	8	US-10-995-561-58398	Sequence 58398, A	C 608	20	0.5	400	6	US-09-925-065A-224389	Sequence 224389,
C 536	20	0.5	201	8	US-10-995-561-59543	Sequence 59543, A	C 609	20	0.5	400	6	US-09-925-065A-224390	Sequence 224390,
C 537	20	0.5	201	8	US-10-995-561-63343	Sequence 63343, A	C 610	20	0.5	414	6	US-09-925-065A-659623	Sequence 659623,
C 538	20	0.5	201	8	US-10-995-561-68821	Sequence 68821, A	C 611	20	0.5	417	8	US-10-775-169-209	Sequence 209, App
C 539	20	0.5	201	8	US-10-995-561-69966	Sequence 69966, A	C 612	20	0.5	419	6	US-09-925-065A-224941	Sequence 224941,
C 540	20	0.5	201	8	US-10-995-561-70143	Sequence 70143, A	C 613	20	0.5	422	6	US-09-925-065A-224942	Sequence 224942,
C 541	20	0.5	201	12	US-11-124-368A-3007	Sequence 3007, A	C 614	20	0.5	422	6	US-09-925-065A-477261	Sequence 477261,
C 542	20	0.5	201	12	US-11-124-368A-3008	Sequence 3008, A	C 615	20	0.5	423	6	US-09-925-065A-580505	Sequence 580505,
C 543	20	0.5	201	12	US-11-124-368A-3032	Sequence 3032, A	C 616	20	0.5	423	6	US-09-925-065A-580506	Sequence 580506,
C 544	20	0.5	201	12	US-11-124-368A-3933	Sequence 3933, A	C 617	20	0.5	423	6	US-09-925-065A-580507	Sequence 580507,
C 545	20	0.5	201	12	US-11-124-368A-5831	Sequence 5831, A	C 618	20	0.5	423	6	US-09-925-065A-580508	Sequence 580508,
C 546	20	0.5	201	12	US-11-124-368A-5832	Sequence 5832, A	C 619	20	0.5	424	6	US-09-925-065A-186016	Sequence 186016,
C 547	20	0.5	201	12	US-11-124-368A-5872	Sequence 5872, A	C 620	20	0.5	424	6	US-09-925-065A-186017	Sequence 186017,
C 548	20	0.5	201	12	US-11-124-368A-5873	Sequence 5873, A	C 621	20	0.5	424	6	US-09-925-065A-186018	Sequence 186018,
C 549	20	0.5	201	12	US-11-124-368A-6015	Sequence 6015, A	C 622	20	0.5	424	6	US-09-925-065A-186019	Sequence 186019,
C 550	20	0.5	201	12	US-11-124-368A-6016	Sequence 6016, A	C 623	20	0.5	429	6	US-09-925-065A-577307	Sequence 577307,
C 551	20	0.5	201	12	US-11-124-368A-9444	Sequence 9444, A	C 624	20	0.5	432	6	US-09-925-065A-173179	Sequence 173179,
C 552	20	0.5	201	12	US-11-124-368A-11180	Sequence 11180, A	C 625	20	0.5	434	6	US-09-925-065A-191228	Sequence 191228,
C 553	20	0.5	201	12	US-11-124-368A-11181	Sequence 11181, A	C 626	20	0.5	435	6	US-09-925-065A-325438	Sequence 325438,
C 554	20	0.5	201	12	US-11-124-368A-11274	Sequence 11274, A	C 627	20	0.5	435	6	US-09-925-065A-168831	Sequence 168831,
C 555	20	0.5	201	12	US-11-124-368A-11275	Sequence 11275, A	C 628	20	0.5	439	6	US-09-925-065A-168832	Sequence 168832,
C 556	20	0.5	201	12	US-11-124-368A-11451	Sequence 11451, A	C 629	20	0.5	439	6	US-09-925-065A-561020	Sequence 561020,
C 557	20	0.5	201	12	US-11-124-368A-11452	Sequence 11452, A	C 630	20	0.5	439	8	US-10-821-234-374	Sequence 374, App
C 558	20	0.5	201	12	US-11-124-368A-12335	Sequence 12335, A	C 631	20	0.5	440	6	US-09-925-065A-393555	Sequence 393555,
C 559	20	0.5	201	12	US-11-124-368A-12336	Sequence 12336, A	C 632	20	0.5	442	6	US-09-925-065A-418222	Sequence 418222,
C 560	20	0.5	201	12	US-11-124-368A-12336	Sequence 12336, A	C 633	20	0.5	446	6	US-09-925-065A-445201	Sequence 445201,
C 561	20	0.5	201	12	US-11-124-368A-14951	Sequence 14951, A	C 634	20	0.5	446	6	US-09-925-065A-445202	Sequence 445202,
C 562	20	0.5	201	12	US-11-124-368A-16389	Sequence 16389, A	C 635	20	0.5	446	6	US-09-925-065A-142670	Sequence 142670,
C 563	20	0.5	201	12	US-11-124-368A-16390	Sequence 16390, A	C 636	20	0.5	452	6	US-09-925-065A-814223	Sequence 814223,
C 564	20	0.5	201	12	US-11-124-368A-18231	Sequence 18231, A	C 637	20	0.5	452	6	US-09-925-065A-111854	Sequence 111854,
C 565	20	0.5	201	12	US-11-124-368A-19762	Sequence 19762, A	C 638	20	0.5	455	6	US-09-925-065A-876285	Sequence 876285,
C 566	20	0.5	201	12	US-11-124-367A-7142	Sequence 7142, A	C 639	20	0.5	473	6	US-09-925-065A-274464	Sequence 274464,
C 567	20	0.5	201	12	US-11-124-367A-8503	Sequence 8503, A	C 640	20	0.5	475	6	US-09-925-065A-455964	Sequence 455964,
C 568	20	0.5	201	12	US-11-124-367A-8540	Sequence 8540, A	C 641	20	0.5	475	6	US-09-925-065A-519059	Sequence 519059,
C 569	20	0.5	201	12	US-11-124-367A-9039	Sequence 9039, A	C 642	20	0.5	475	6	US-09-925-065A-884545	Sequence 884545,
C 570	20	0.5	201	12	US-11-124-367A-9349	Sequence 9349, A	C 643	20	0.5	476	6	US-09-925-065A-295118	Sequence 295118,
C 571	20	0.5	201	12	US-11-124-367A-9937	Sequence 9937, A	C 644	20	0.5	478	6	US-09-925-065A-787904	Sequence 787904,
C 572	20	0.5	201	12	US-11-124-367A-9938	Sequence 9938, A	C 645	20	0.5	480	6	US-09-925-065A-565124	Sequence 565124,
C 573	20	0.5	201	12	US-11-124-367A-10555	Sequence 10555, A	C 646	20	0.5	481	6	US-09-925-065A-239505	Sequence 239505,
C 574	20	0.5	201	12	US-11-124-367A-10556	Sequence 10556, A	C 647	20	0.5	481	6	US-09-925-065A-239506	Sequence 239506,
C 575	20	0.5	201	12	US-11-124-367A-14032	Sequence 14032, A	C 648	20	0.5	481	6	US-09-925-065A-277231	Sequence 277231,
C 576	20	0.5	201	12	US-11-124-367A-15496	Sequence 15496, A	C 649	20	0.5	481	6	US-09-925-065A-616435	Sequence 616435,
C 577	20	0.5	201	12	US-11-124-367A-20469	Sequence 20469, A	C 650	20	0.5	481	6	US-09-925-065A-616436	Sequence 616436,
C 578	20	0.5	201	12	US-11-124-367A-20895	Sequence 20895, A	C 651	20	0.5	482	6	US-09-925-065A-49808	Sequence 49808, A
C 579	20	0.5	201	12	US-11-124-367A-20901	Sequence 20901, A	C 652	20	0.5	482	6	US-09-925-065A-205973	Sequence 205973,
C 580	20	0.5	201	12	US-11-124-367A-21408	Sequence 21408, A	C 653	20	0.5	483	6	US-09-925-065A-184098	Sequence 184098,
C 581	20	0.5	201	12	US-11-124-367A-22986	Sequence 22986, A	C 654	20	0.5	483	6	US-09-925-065A-420033	Sequence 420033,
C 582	20	0.5	201	12	US-11-124-367A-22987	Sequence 22987, A	C 655	20	0.5	484	6	US-09-925-065A-249720	Sequence 249720,
C 583	20	0.5	201	12	US-11-124-367A-24501	Sequence 24501, A	C 656	20	0.5	489	6	US-09-925-065A-642169	Sequence 642169,
C 584	20	0.5	201	12	US-11-124-367A-27827	Sequence 27827, A	C 657	20	0.5	489	6	US-09-925-065A-97802	Sequence 97802, A
C 585	20	0.5	201	12	US-11-124-367A-27965	Sequence 27965, A	C 658	20	0.5	490	6	US-09-925-065A-97803	Sequence 97803, A
C 586	20	0.5	201	12	US-11-124-367A-28174	Sequence 28174, A	C 659	20	0.5	493	6	US-09-925-065A-519299	Sequence 519299,
C 587	20	0.5	201	12	US-11-124-367A-28212	Sequence 28212, A	C 660	20	0.5	493	6	US-09-925-065A-519300	Sequence 519300,
C 588	20	0.5	201	12	US-11-124-367A-29432	Sequence 29432, A	C 661	20	0.5	494	6	US-09-925-065A-569095	Sequence 569095,
C 589	20	0.5	201	12	US-11-124-367A-29491	Sequence 29491, A	C 662	20	0.5	494	6	US-09-925-065A-569096	Sequence 569096,
C 590	20	0.5	201	12	US-11-124-367A-29596	Sequence 29596, A	C 663	20	0.5	495	6	US-09-925-065A-173150	Sequence 173150,
C 591	20	0.5	201	12	US-11-124-367A-32573	Sequence 32573, A	C 664	20	0.5	495	6	US-09-925-065A-173152	Sequence 173152,
C 592	20	0.5	201	12	US-11-124-367A-33933	Sequence 33933, A	C 665	20	0.5	495	6	US-09-925-065A-173153	Sequence 173153,
C 593	20	0.5	209	6	US-09-925-065A-196854	Sequence 196854, A	C 666	20	0.5	495	6	US-09-925-065A-173154	Sequence 173154,
C 594	20	0.5	209	6	US-09-925-065A-196855	Sequence 196855, A	C 667	20	0.5	497	6	US-09-925-065A-9781	Sequence 9781, Ap
C 595	20	0.5	281	6	US-09-925-065A-311771	Sequence 311771, A	C 668	20	0.5	498	6	US-09-925-065A-413104	Sequence 413104,
C 596	20	0.5	281	6	US-09-925-065A-311772	Sequence 311772, A	C 669	20	0.5	501	6	US-09-925-065A-28311	Sequence 28311, A
C 597	20	0.5	281	6	US-09-925-065A-311773	Sequence 311773, A	C 670	20	0.5	501	6	US-09-925-065A-158135	Sequence 158135,
C 598	20	0.5	281	6	US-09-925-065A-311774	Sequence 311774, A	C 671	20	0.5	501	6	US-09-925-065A-161612	Sequence 161612,
C 599	20	0.5	309	6	US-09-925-065A-239880	Sequence 239880, A	C 672	20	0.5	503	6	US-09-925-065A-664563	Sequence 664563,
C 600	20	0.5	318	6	US-09-925-065A-452194	Sequence 452194, A	C 673	20	0.5	505	6	US-09-925-065A-694214	Sequence 694214,
C 601	20	0.5	330	6	US-09-925-065A-452194	Sequence 452194, A	C 674	20	0.5	506	6	US-09-925-065A-824298	Sequence 824298,
C 602	20	0.5	330	8	US-10-508-424-3	Sequence 145591, A	C 675	20	0.5	507	6	US-09-925-065A-507282	Sequence 507282,
C 603	20	0.5	352	6	US-09-925-065A-605018	Sequence 3, Appli	C 676	20	0.5	508	6	US-09-925-065A-13624	Sequence 13624, A
C 604	20	0.5	352	6	US-09-925-065A-605019	Sequence 605019, A	C 677	20	0.5	510	6	US-09-925-065A-199354	Sequence 199354,



678	20	0.5	510	6	US-09-925-065A-199355	Sequence 199355,	c 751	20	0.5	547	6	US-09-925-065A-335086	Sequence 335086,
679	20	0.5	510	6	US-09-925-065A-199356	Sequence 199356,	c 752	20	0.5	547	6	US-09-925-065A-335087	Sequence 335087,
680	20	0.5	510	6	US-09-925-065A-199357	Sequence 199357,	c 753	20	0.5	548	6	US-09-925-065A-141108	Sequence 141108,
c 681	20	0.5	512	6	US-09-925-065A-801159	Sequence 801159,	754	20	0.5	548	6	US-09-925-065A-141109	Sequence 141109,
682	20	0.5	512	6	US-09-925-065A-837112	Sequence 837112,	755	20	0.5	548	6	US-09-925-065A-141110	Sequence 141110,
c 683	20	0.5	516	6	US-09-925-065A-185890	Sequence 185890,	756	20	0.5	548	6	US-09-925-065A-932672	Sequence 932672,
c 684	20	0.5	516	6	US-09-925-065A-185891	Sequence 185891,	c 757	20	0.5	548	6	US-09-925-065A-940576	Sequence 940576,
685	20	0.5	516	6	US-09-925-065A-622900	Sequence 622900,	758	20	0.5	549	6	US-09-925-065A-584164	Sequence 584164,
686	20	0.5	517	6	US-09-925-065A-596198	Sequence 596198,	759	20	0.5	549	6	US-09-925-065A-584165	Sequence 584165,
c 687	20	0.5	519	6	US-09-925-065A-435920	Sequence 435920,	c 760	20	0.5	550	6	US-09-925-065A-226093	Sequence 226093,
c 688	20	0.5	520	6	US-09-925-065A-42513	Sequence 42513, A	761	20	0.5	550	6	US-09-925-065A-415667	Sequence 415667,
c 689	20	0.5	520	6	US-09-925-065A-651813	Sequence 651813,	762	20	0.5	550	6	US-09-925-065A-917250	Sequence 917250,
c 690	20	0.5	521	6	US-09-925-065A-100028	Sequence 100028,	763	20	0.5	550	6	US-09-925-065A-917251	Sequence 917251,
c 691	20	0.5	521	6	US-09-925-065A-207449	Sequence 207449,	764	20	0.5	550	6	US-09-925-065A-944580	Sequence 944580,
692	20	0.5	521	6	US-09-925-065A-475183	Sequence 475183,	765	20	0.5	552	6	US-09-925-065A-336555	Sequence 336555,
c 693	20	0.5	522	6	US-09-925-065A-435139	Sequence 435139,	c 766	20	0.5	552	6	US-09-925-065A-364387	Sequence 364387,
c 694	20	0.5	522	6	US-09-925-065A-435140	Sequence 435140,	767	20	0.5	552	6	US-09-925-065A-784825	Sequence 784825,
c 695	20	0.5	522	6	US-09-925-065A-751095	Sequence 751095,	768	20	0.5	552	6	US-09-925-065A-794432	Sequence 794432,
696	20	0.5	523	6	US-09-925-065A-330399	Sequence 330399,	769	20	0.5	554	6	US-09-925-065A-312374	Sequence 312374,
697	20	0.5	523	6	US-09-925-065A-330400	Sequence 330400,	770	20	0.5	554	6	US-09-925-065A-312375	Sequence 312375,
698	20	0.5	523	6	US-09-925-065A-330401	Sequence 330401,	771	20	0.5	555	6	US-09-925-065A-886818	Sequence 886818,
c 699	20	0.5	523	6	US-09-925-065A-901309	Sequence 901309,	772	20	0.5	556	6	US-09-925-065A-20152	Sequence 20152, A
c 700	20	0.5	524	6	US-09-925-065A-151888	Sequence 151888,	773	20	0.5	556	6	US-09-925-065A-481201	Sequence 481201,
c 701	20	0.5	524	6	US-09-925-065A-164525	Sequence 164525,	774	20	0.5	556	6	US-09-925-065A-481202	Sequence 481202,
c 702	20	0.5	524	6	US-09-925-065A-514873	Sequence 514873,	775	20	0.5	556	6	US-09-925-065A-481203	Sequence 481203,
703	20	0.5	525	6	US-09-925-065A-322765	Sequence 322765,	776	20	0.5	556	6	US-09-925-065A-481204	Sequence 481204,
704	20	0.5	525	6	US-09-925-065A-564079	Sequence 564079,	c 777	20	0.5	556	6	US-09-925-065A-607342	Sequence 607342,
c 705	20	0.5	526	6	US-09-925-065A-340981	Sequence 340981,	778	20	0.5	557	6	US-09-925-065A-732510	Sequence 732510,
c 706	20	0.5	528	6	US-09-925-065A-193474	Sequence 193474,	c 779	20	0.5	557	6	US-09-925-065A-782216	Sequence 782216,
707	20	0.5	528	6	US-09-925-065A-493300	Sequence 493300,	780	20	0.5	558	6	US-09-925-065A-911540	Sequence 911540,
c 708	20	0.5	528	6	US-09-925-065A-645698	Sequence 645698,	c 781	20	0.5	559	6	US-09-925-065A-161089	Sequence 161089,
c 709	20	0.5	528	6	US-09-925-065A-645699	Sequence 645699,	c 782	20	0.5	559	6	US-09-925-065A-161090	Sequence 161090,
c 710	20	0.5	528	6	US-09-925-065A-645700	Sequence 645700,	c 783	20	0.5	559	6	US-09-925-065A-324822	Sequence 324822,
c 711	20	0.5	530	6	US-09-925-065A-37062	Sequence 37062, A	c 784	20	0.5	559	6	US-09-925-065A-786982	Sequence 786982,
c 712	20	0.5	530	6	US-09-925-065A-37063	Sequence 37063, A	785	20	0.5	560	6	US-09-925-065A-439872	Sequence 439872,
c 713	20	0.5	530	6	US-09-925-065A-37064	Sequence 37064, A	786	20	0.5	560	6	US-09-925-065A-439873	Sequence 439873,
c 714	20	0.5	530	6	US-09-925-065A-37065	Sequence 37065, A	787	20	0.5	560	6	US-09-925-065A-439874	Sequence 439874,
c 715	20	0.5	530	6	US-09-925-065A-37066	Sequence 37066, A	c 788	20	0.5	561	6	US-09-925-065A-16247	Sequence 16247, A
c 716	20	0.5	532	6	US-09-925-065A-204619	Sequence 204619,	c 789	20	0.5	561	6	US-09-925-065A-141833	Sequence 141833,
c 717	20	0.5	534	6	US-09-925-065A-386339	Sequence 386339,	c 790	20	0.5	561	6	US-09-925-065A-193366	Sequence 193366,
c 718	20	0.5	534	6	US-09-925-065A-474669	Sequence 474669,	c 791	20	0.5	563	6	US-09-925-065A-348017	Sequence 348017,
719	20	0.5	534	6	US-09-925-065A-597928	Sequence 597928,	c 792	20	0.5	564	6	US-09-925-065A-20066	Sequence 20066, A
c 720	20	0.5	535	6	US-09-925-065A-597929	Sequence 597929,	c 793	20	0.5	564	6	US-09-925-065A-219003	Sequence 219003,
c 721	20	0.5	535	6	US-09-925-065A-417526	Sequence 417526,	c 794	20	0.5	564	6	US-09-925-065A-294101	Sequence 294101,
c 722	20	0.5	537	6	US-09-925-065A-382126	Sequence 382126,	795	20	0.5	564	6	US-09-925-065A-294102	Sequence 294102,
c 723	20	0.5	537	6	US-09-925-065A-382127	Sequence 382127,	796	20	0.5	564	6	US-09-925-065A-294103	Sequence 294103,
c 724	20	0.5	538	6	US-09-925-065A-346150	Sequence 346150,	797	20	0.5	564	6	US-09-925-065A-477681	Sequence 477681,
725	20	0.5	539	6	US-09-925-065A-244554	Sequence 244554,	c 798	20	0.5	564	6	US-09-925-065A-801414	Sequence 801414,
c 726	20	0.5	539	6	US-09-925-065A-244555	Sequence 244555,	c 799	20	0.5	566	6	US-09-925-065A-569811	Sequence 569811,
c 727	20	0.5	540	6	US-09-925-065A-157112	Sequence 157112,	c 800	20	0.5	565	6	US-09-925-065A-569812	Sequence 569812,
c 728	20	0.5	540	6	US-09-925-065A-206002	Sequence 206002,	c 801	20	0.5	565	6	US-09-925-065A-569812	Sequence 569812,
729	20	0.5	540	6	US-09-925-065A-260994	Sequence 260994,	c 802	20	0.5	565	6	US-09-925-065A-737031	Sequence 737031,
c 730	20	0.5	540	6	US-09-925-065A-260995	Sequence 260995,	c 803	20	0.5	565	6	US-09-925-065A-810062	Sequence 810062,
c 731	20	0.5	540	6	US-09-925-065A-316436	Sequence 316436,	c 804	20	0.5	566	6	US-09-925-065A-144930	Sequence 144930,
732	20	0.5	541	6	US-09-925-065A-872131	Sequence 872131,	c 805	20	0.5	566	6	US-09-925-065A-839288	Sequence 839288,
733	20	0.5	541	6	US-09-925-065A-872131	Sequence 872131,	c 806	20	0.5	567	6	US-09-925-065A-249887	Sequence 249887,
c 734	20	0.5	541	6	US-09-925-065A-902810	Sequence 902810,	c 807	20	0.5	567	6	US-09-925-065A-542764	Sequence 542764,
c 735	20	0.5	541	6	US-09-925-065A-909125	Sequence 909125,	c 808	20	0.5	568	6	US-09-925-065A-268921	Sequence 268921,
c 736	20	0.5	542	6	US-09-925-065A-106382	Sequence 106382,	c 809	20	0.5	568	6	US-09-925-065A-268922	Sequence 268922,
c 737	20	0.5	542	6	US-09-925-065A-403114	Sequence 403114,	c 810	20	0.5	570	6	US-09-925-065A-399162	Sequence 399162,
c 738	20	0.5	542	6	US-09-925-065A-403115	Sequence 403115,	c 811	20	0.5	570	6	US-09-925-065A-403113	Sequence 403113,
739	20	0.5	542	6	US-09-925-065A-777562	Sequence 777562,	c 812	20	0.5	570	6	US-09-925-065A-507552	Sequence 507552,
c 740	20	0.5	543	6	US-09-925-065A-743551	Sequence 743551,	c 813	20	0.5	570	6	US-09-925-065A-397057	Sequence 397057,
c 741	20	0.5	543	6	US-09-925-065A-840730	Sequence 840730,	c 814	20	0.5	571	6	US-09-925-065A-397058	Sequence 397058,
c 742	20	0.5	543	6	US-09-925-065A-840731	Sequence 840731,	c 815	20	0.5	571	6	US-09-925-065A-602511	Sequence 602511,
743	20	0.5	544	6	US-09-925-065A-51783	Sequence 51783, A	c 816	20	0.5	571	6	US-09-925-065A-619649	Sequence 619649,
744	20	0.5	544	6	US-09-925-065A-179232	Sequence 179232,	c 817	20	0.5	571	6	US-09-925-065A-619650	Sequence 619650,
c 745	20	0.5	544	6	US-09-925-065A-479406	Sequence 479406,	c 818	20	0.5	571	6	US-09-925-065A-619651	Sequence 619651,
c 746	20	0.5	546	6	US-09-925-065A-372299	Sequence 372299, A	c 819	20	0.5	571	6	US-09-925-065A-944418	Sequence 944418,
c 747	20	0.5	546	6	US-09-925-065A-243317	Sequence 243317,	c 820	20	0.5	572	6	US-09-925-065A-957059	Sequence 957059,
c 748	20	0.5	547	6	US-09-925-065A-20800	Sequence 20800, A	c 821	20	0.5	572	6	US-09-925-065A-245295	Sequence 245295,
c 749	20	0.5	547	6	US-09-925-065A-20801	Sequence 20801, A	c 822	20	0.5	572	6	US-09-925-065A-245296	Sequence 245296,
c 750	20	0.5	547	6	US-09-925-065A-335085	Sequence 335085,	c 823	20	0.5	572	6	US-09-925-065A-245297	Sequence 245297,



C 824	20	0.5	573	6	US-09-925-065A-862992	Sequence 862992,	897	20	0.5	597	6	US-09-925-065A-544338	Sequence 544338,
C 825	20	0.5	573	6	US-09-925-065A-881144	Sequence 881144,	C 898	20	0.5	598	6	US-09-925-065A-747157	Sequence 747157,
C 826	20	0.5	574	6	US-09-925-065A-755518	Sequence 755518,	C 899	20	0.5	598	6	US-09-925-065A-813805	Sequence 813805,
C 827	20	0.5	574	6	US-09-925-065A-642047	Sequence 642047,	C 900	20	0.5	599	6	US-09-925-065A-317058	Sequence 317058,
C 828	20	0.5	574	6	US-09-925-065A-785923	Sequence 785923,	C 901	20	0.5	600	6	US-09-925-065A-92089	Sequence 92089, A
C 829	20	0.5	574	6	US-09-925-065A-868726	Sequence 868726,	C 902	20	0.5	600	6	US-09-925-065A-92090	Sequence 92090, A
C 830	20	0.5	575	6	US-09-925-065A-156095	Sequence 156095,	C 903	20	0.5	600	6	US-09-925-065A-723981	Sequence 723981,
C 831	20	0.5	575	6	US-09-925-065A-458505	Sequence 458505,	C 904	20	0.5	600	8	US-10-750-185-3687	Sequence 3687, Ap
C 832	20	0.5	575	6	US-09-925-065A-458506	Sequence 458506,	C 905	20	0.5	600	8	US-10-750-623-3687	Sequence 3687, Ap
C 833	20	0.5	575	6	US-09-925-065A-585880	Sequence 585880,	C 906	20	0.5	600	12	US-11-136-527-5874	Sequence 5874, Ap
C 834	20	0.5	575	6	US-09-925-065A-866034	Sequence 866034,	C 907	20	0.5	601	6	US-09-925-065A-609447	Sequence 609447,
C 835	20	0.5	576	6	US-09-925-065A-554742	Sequence 554742,	C 908	20	0.5	601	6	US-09-925-065A-886614	Sequence 886614,
C 836	20	0.5	576	6	US-09-925-065A-554743	Sequence 554743,	C 909	20	0.5	601	6	US-09-925-065A-905649	Sequence 905649,
C 837	20	0.5	576	6	US-09-925-065A-587473	Sequence 587473,	C 910	20	0.5	601	6	US-09-925-065A-951355	Sequence 951355,
C 838	20	0.5	576	6	US-09-925-065A-587473	Sequence 587473,	C 911	20	0.5	602	6	US-09-925-065A-728588	Sequence 728588,
C 839	20	0.5	576	6	US-09-925-065A-773959	Sequence 773959,	C 912	20	0.5	602	6	US-09-925-065A-951355	Sequence 951355,
C 840	20	0.5	576	6	US-09-925-065A-837083	Sequence 837083,	C 913	20	0.5	602	6	US-09-925-065A-728589	Sequence 728589,
C 841	20	0.5	577	6	US-09-925-065A-568812	Sequence 568812,	C 914	20	0.5	602	6	US-09-925-065A-728590	Sequence 728590,
C 842	20	0.5	577	6	US-09-925-065A-768564	Sequence 768564,	C 915	20	0.5	603	6	US-09-925-065A-745421	Sequence 745421,
C 843	20	0.5	577	6	US-09-925-065A-847406	Sequence 847406,	C 916	20	0.5	603	6	US-09-925-065A-349601	Sequence 349601,
C 844	20	0.5	577	6	US-09-925-065A-905486	Sequence 905486,	C 917	20	0.5	603	6	US-09-925-065A-349602	Sequence 349602,
C 845	20	0.5	578	6	US-09-925-065A-869771	Sequence 869771,	C 918	20	0.5	604	6	US-09-925-065A-188736	Sequence 188736,
C 846	20	0.5	579	6	US-09-925-065A-27920	Sequence 27920, A	C 919	20	0.5	604	6	US-09-925-065A-188732	Sequence 188732,
C 847	20	0.5	579	6	US-09-925-065A-168373	Sequence 168373,	C 920	20	0.5	604	6	US-09-925-065A-764789	Sequence 764789,
C 848	20	0.5	580	6	US-09-925-065A-433786	Sequence 433786,	C 921	20	0.5	605	6	US-09-925-065A-570156	Sequence 570156,
C 849	20	0.5	580	6	US-09-925-065A-480558	Sequence 480558,	C 922	20	0.5	605	6	US-09-925-065A-572277	Sequence 572277,
C 850	20	0.5	580	6	US-09-925-065A-480559	Sequence 480559,	C 923	20	0.5	605	6	US-09-925-065A-764788	Sequence 764788,
C 851	20	0.5	580	6	US-09-925-065A-480560	Sequence 480560,	C 924	20	0.5	606	6	US-09-925-065A-529361	Sequence 529361,
C 852	20	0.5	581	6	US-09-925-065A-113197	Sequence 113197,	C 925	20	0.5	606	6	US-09-925-065A-907929	Sequence 907929,
C 853	20	0.5	581	6	US-09-925-065A-379472	Sequence 379472,	C 926	20	0.5	607	6	US-09-925-065A-405550	Sequence 405550,
C 854	20	0.5	581	6	US-09-925-065A-642170	Sequence 642170,	C 927	20	0.5	607	6	US-09-925-065A-405551	Sequence 405551,
C 855	20	0.5	581	6	US-09-925-065A-738893	Sequence 738893,	C 928	20	0.5	607	6	US-09-925-065A-754660	Sequence 754660,
C 856	20	0.5	581	6	US-09-925-065A-815514	Sequence 815514,	C 929	20	0.5	607	6	US-09-925-065A-825392	Sequence 825392,
C 857	20	0.5	582	6	US-09-925-065A-867990	Sequence 867990,	C 930	20	0.5	608	6	US-09-925-065A-85313	Sequence 85313, A
C 858	20	0.5	582	6	US-09-925-065A-875488	Sequence 875488,	C 931	20	0.5	608	6	US-09-925-065A-85314	Sequence 85314, A
C 859	20	0.5	582	6	US-09-925-065A-882716	Sequence 882716,	C 932	20	0.5	608	6	US-09-925-065A-85315	Sequence 85315, A
C 860	20	0.5	583	6	US-09-925-065A-336092	Sequence 336092,	C 933	20	0.5	608	6	US-09-925-065A-85316	Sequence 85316, A
C 861	20	0.5	583	6	US-09-925-065A-922112	Sequence 922112,	C 934	20	0.5	608	6	US-09-925-065A-85317	Sequence 85317, A
C 862	20	0.5	583	6	US-09-925-065A-946852	Sequence 946852,	C 935	20	0.5	608	6	US-09-925-065A-321867	Sequence 321867,
C 863	20	0.5	583	6	US-09-925-065A-946853	Sequence 946853,	C 936	20	0.5	608	6	US-09-925-065A-352180	Sequence 352180,
C 864	20	0.5	584	6	US-09-925-065A-831357	Sequence 831357,	C 937	20	0.5	609	6	US-09-925-065A-702270	Sequence 702270,
C 865	20	0.5	585	6	US-09-925-065A-164356	Sequence 164356,	C 938	20	0.5	609	6	US-09-925-065A-732103	Sequence 732103,
C 866	20	0.5	585	6	US-09-925-065A-164357	Sequence 164357,	C 939	20	0.5	610	6	US-09-925-065A-443704	Sequence 443704,
C 867	20	0.5	585	6	US-09-925-065A-190730	Sequence 190730,	C 940	20	0.5	610	6	US-09-925-065A-508515	Sequence 508515,
C 868	20	0.5	585	6	US-09-925-065A-477443	Sequence 477443,	C 941	20	0.5	610	6	US-09-925-065A-917540	Sequence 917540,
C 869	20	0.5	585	6	US-09-925-065A-506226	Sequence 506226,	C 942	20	0.5	611	6	US-09-925-065A-797466	Sequence 797466,
C 870	20	0.5	585	6	US-09-925-065A-580233	Sequence 580233,	C 943	20	0.5	611	6	US-09-925-065A-842985	Sequence 842985,
C 871	20	0.5	585	6	US-09-925-065A-841932	Sequence 841932,	C 944	20	0.5	611	6	US-09-925-065A-842985	Sequence 842985,
C 872	20	0.5	587	6	US-09-925-065A-144688	Sequence 144688,	C 945	20	0.5	612	6	US-09-925-065A-504072	Sequence 504072,
C 873	20	0.5	587	6	US-09-925-065A-144689	Sequence 144689,	C 946	20	0.5	612	6	US-09-925-065A-783786	Sequence 783786,
C 874	20	0.5	587	6	US-09-925-065A-742471	Sequence 742471,	C 947	20	0.5	612	6	US-09-925-065A-938133	Sequence 938133,
C 875	20	0.5	588	6	US-09-925-065A-729746	Sequence 729746,	C 948	20	0.5	612	6	US-09-925-065A-938134	Sequence 938134,
C 876	20	0.5	588	6	US-09-925-065A-774184	Sequence 774184,	C 949	20	0.5	613	6	US-09-925-065A-107683	Sequence 107683,
C 877	20	0.5	589	6	US-09-925-065A-118525	Sequence 118525,	C 950	20	0.5	614	6	US-09-925-065A-743489	Sequence 743489,
C 878	20	0.5	589	6	US-09-925-065A-906774	Sequence 906774,	C 951	20	0.5	615	6	US-09-925-065A-516141	Sequence 516141,
C 879	20	0.5	590	6	US-09-925-065A-596567	Sequence 596567,	C 952	20	0.5	616	6	US-09-925-065A-787828	Sequence 787828,
C 880	20	0.5	590	6	US-09-925-065A-596568	Sequence 596568,	C 953	20	0.5	616	6	US-09-925-065A-897859	Sequence 897859,
C 881	20	0.5	591	6	US-09-925-065A-631554	Sequence 631554,	C 954	20	0.5	616	6	US-09-925-065A-110476	Sequence 110476,
C 882	20	0.5	591	6	US-09-925-065A-631555	Sequence 631555,	C 955	20	0.5	619	6	US-09-925-065A-110477	Sequence 110477,
C 883	20	0.5	591	6	US-09-925-065A-631556	Sequence 631556,	C 956	20	0.5	619	6	US-09-925-065A-232244	Sequence 232244,
C 884	20	0.5	591	6	US-09-925-065A-631557	Sequence 631557,	C 957	20	0.5	619	6	US-09-925-065A-761988	Sequence 761988,
C 885	20	0.5	592	6	US-09-925-065A-263345	Sequence 263345,	C 958	20	0.5	619	6	US-09-925-065A-761989	Sequence 761989,
C 886	20	0.5	592	6	US-09-925-065A-785378	Sequence 785378,	C 959	20	0.5	619	6	US-09-925-065A-774097	Sequence 774097,
C 887	20	0.5	592	6	US-09-925-065A-660686	Sequence 660686,	C 960	20	0.5	619	6	US-09-925-065A-830263	Sequence 830263,
C 888	20	0.5	594	6	US-09-925-065A-12141	Sequence 12141, A	C 961	20	0.5	619	6	US-09-925-065A-830264	Sequence 830264,
C 889	20	0.5	594	6	US-09-925-065A-607049	Sequence 607049,	C 962	20	0.5	620	6	US-09-925-065A-830264	Sequence 830264,
C 890	20	0.5	594	6	US-09-925-065A-690201	Sequence 690201,	C 963	20	0.5	620	6	US-09-925-065A-832218	Sequence 832218,
C 891	20	0.5	595	6	US-09-925-065A-259255	Sequence 259255,	C 964	20	0.5	621	6	US-09-925-065A-798426	Sequence 798426,
C 892	20	0.5	595	6	US-09-925-065A-882180	Sequence 882180,	C 965	20	0.5	623	6	US-09-925-065A-725778	Sequence 725778,
C 893	20	0.5	595	6	US-09-925-065A-882181	Sequence 882181,	C 966	20	0.5	625	6	US-09-925-065A-242563	Sequence 242563,
C 894	20	0.5	596	6	US-09-925-065A-187708	Sequence 187708,	C 967	20	0.5	625	6	US-09-925-065A-670349	Sequence 670349,
C 895	20	0.5	596	6	US-09-925-065A-333100	Sequence 333100,	C 968	20	0.5	625	6	US-09-925-065A-705261	Sequence 705261,
C 896	20	0.5	596	6	US-09-925-065A-583870	Sequence 583870,	C 969	20	0.5	625	6	US-09-925-065A-705262	Sequence 705262,

970 20 0.5 625 6 US-09-925-065A-874597, Sequence 874597,  
c 971 20 0.5 626 6 US-09-925-065A-716165, Sequence 716165,  
c 972 20 0.5 626 6 US-09-925-065A-716166, Sequence 716166,  
c 973 20 0.5 626 6 US-09-925-065A-716167, Sequence 716167,  
c 974 20 0.5 626 6 US-09-925-065A-716168, Sequence 716168,  
c 975 20 0.5 626 6 US-09-925-065A-716169, Sequence 716169,  
c 976 20 0.5 629 6 US-09-925-065A-544464, Sequence 544464,  
c 977 20 0.5 631 6 US-09-925-065A-230271, Sequence 230271,  
c 978 20 0.5 631 6 US-09-925-065A-37838, Sequence 37838, A  
c 979 20 0.5 631 6 US-09-925-065A-37839, Sequence 37839, A  
c 980 20 0.5 631 6 US-09-925-065A-37840, Sequence 37840, A  
c 981 20 0.5 632 6 US-09-925-065A-386701, Sequence 386701,  
c 982 20 0.5 632 6 US-09-925-065A-677057, Sequence 677057,  
c 983 20 0.5 632 6 US-09-925-065A-677058, Sequence 677058,  
c 984 20 0.5 632 6 US-09-925-065A-773381, Sequence 773381,  
c 985 20 0.5 632 6 US-09-925-065A-900492, Sequence 900492,  
c 986 20 0.5 634 6 US-09-925-065A-318173, Sequence 318173,  
c 987 20 0.5 636 6 US-09-925-065A-745004, Sequence 745004,  
c 988 20 0.5 636 6 US-09-925-065A-777105, Sequence 777105,  
c 989 20 0.5 636 6 US-10-973-115B-507, Sequence 507, App  
c 990 20 0.5 637 6 US-09-925-065A-948087, Sequence 948087,  
c 991 20 0.5 637 6 US-09-925-065A-956472, Sequence 956472,  
c 992 20 0.5 638 6 US-09-925-065A-727734, Sequence 727734,  
c 993 20 0.5 638 6 US-09-925-065A-759540, Sequence 759540,  
c 994 20 0.5 639 6 US-09-925-065A-506407, Sequence 506407,  
c 995 20 0.5 639 6 US-09-925-065A-743657, Sequence 743657,  
c 996 20 0.5 639 6 US-09-925-065A-787827, Sequence 787827,  
c 997 20 0.5 640 6 US-09-925-065A-747816, Sequence 747816,  
c 998 20 0.5 640 6 US-09-925-065A-821187, Sequence 821187,  
c 999 20 0.5 641 6 US-09-925-065A-935392, Sequence 935392,  
1000 20 0.5 642 6 US-09-925-065A-72710, Sequence 72710, A

## ALIGNMENTS

RESULT 1  
US-09-925-065A-727226/c  
; Sequence 727226, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 727226  
; LENGTH: 1275  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-727226

Query Match 3.2%; Score 124; DB 6; Length 1275;  
Best Local Similarity 100.0%; Pred. No. 1.5e-29;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1893 GATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATCCATAGGAATCTTCGAGCT 1952  
DB 543 GATAGTGAATGCTGGAACCTCCCTGTGTATGATGATATCCATAGGAATCTTCGAGCT 484

QY 1953 CTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGC 2012  
DB 483 CTGTGAAGATCTCATCTGGAATAAAGACCCCTGAGGCCACTGAGAAGCTCTTACGTTATGC 424

QY 2013 CCAG 2016  
DB 423 CCAG 420

## RESULT 2

US-09-925-065A-697175/c  
; Sequence 697175, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 697175  
; LENGTH: 616  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-697175

Query Match 2.0%; Score 80; DB 6; Length 616;  
Best Local Similarity 100.0%; Pred. No. 8.2e-16;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2739 GTGTTCCAGCTGTAGTGAATAAATCTAAAGGATGAATATCTTGGAGGAATCATGGAAGA 2798  
DB 134 GTGTTCCAGCTGTAGTGAATAAATCTAAAGGATGAATATCTTGGAGGAATCATGGAAGA 75  
QY 2799 ATATGAAGATATTAGACAGG 2818  
DB 74 ATATGAAGATATTAGACAGG 55

## RESULT 3

US-09-925-065A-460077/c  
; Sequence 460077, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 460077
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-460077

Query Match      1.0%; Score 39; DB 6; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 989 AGGATTTTGGTATGATGGCTTGTCAATATAGTTGGAG 1027
Db |||||||
39 AGGATTTTGGTATGATGGCTTGTCAATATAGTTGGAG 1

RESULT 4
US-09-925-065A-39453/c
; Sequence 39453, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39453
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-39453

Query Match      0.8%; Score 31; DB 6; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3889 TCTTGATGATCCTCAAGGAATACACCTAG 3919
Db |||||||
570 TCTTGATGATCCTCAAGGAATACACCTAG 540

RESULT 5
US-09-925-065A-602656
; Sequence 602656, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 602656
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-602656

Query Match      0.7%; Score 26; DB 6; Length 559;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3864 TTTTGTGCTTTTGGCTTTTAT 3889
Db |||||||
21 TTTTGTGCTTTTGGCTTTTAT 46

RESULT 6
US-09-925-065A-271845
; Sequence 271845, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 271845
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271845

Query Match      0.7%; Score 26; DB 6; Length 595;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTTGGCTTTTAT 3888
Db |||||||
378 TTTTGTGCTTTTGGCTTTTAT 403

RESULT 7
US-09-925-065A-702090
; Sequence 702090, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
```

```
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 702090
; LENGTH: 1374
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-702090

Query Match      0.7%; Score 26; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3862 CTTTTTTTTTTTTTTTGCCTTTTTT 3887
Db 204 CTTTTTTTTTTTTTTTGCCTTTTTT 229

RESULT 8
US-11-175-859-4373
; Sequence 4373, Application US/11175859
; Publication No. US20060024715A1
; GENERAL INFORMATION:
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Method of Analysis of Human Polymorphism
; FILE REFERENCE: 3690.1
; CURRENT APPLICATION NUMBER: US/11/175,859
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: US 60/585,352
; PRIOR FILING DATE: 2004-07-02
; NUMBER OF SEQ ID NOS: 116251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4373
; LENGTH: 50
; TYPE: DNA
; ORGANISM: homo sapien
US-11-175-859-4373

Query Match      0.6%; Score 25; DB 12; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2794 GAAGAATATGAAGATATTAGACAGG 2818
Db 1 GAAGAATATGAAGATATTAGACAGG 25

RESULT 9
US-09-925-065A-279223/c
; Sequence 279223, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
```

```
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 279223
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-279223

Query Match      0.6%; Score 25; DB 6; Length 542;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTTTTTTTTTTTTGG 3879
Db 435 AGACTAACTTTTTTTTTTTTTTTTGG 411

RESULT 10
US-09-925-065A-936651/c
; Sequence 936651, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 936651
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-936651

Query Match      0.6%; Score 25; DB 6; Length 545;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTTTTTTTTTTTTGCCTTTTTT 3887
Db 157 TTTTTTTTTTTTTTTTGCCTTTTTT 133

RESULT 11
US-09-925-065A-550047/c
; Sequence 550047, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
```

; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 550047  
; LENGTH: 630  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-550047

Query Match 0.6%; Score 25; DB 6; Length 630;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTTTTTTTTTTTTGG 3879  
|||||  
Db 582 AGACTAACTTTTTTTTTTTTTTTG 558

## RESULT 12

US-09-925-065A-34355  
; Sequence 34355, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34355  
; LENGTH: 1660  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-34355

Query Match 0.6%; Score 25; DB 6; Length 1660;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTTTTTTTTTTTG 3879  
|||||  
Db 692 AGACTAACTTTTTTTTTTTTTTTG 716

## RESULT 13

US-09-925-065A-34356  
; Sequence 34356, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34356  
; LENGTH: 1660  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-34356

Query Match 0.6%; Score 25; DB 6; Length 1660;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTTTTTTTTTTTG 3879  
|||||  
Db 692 AGACTAACTTTTTTTTTTTTTTTG 716

## RESULT 14

US-09-925-065A-34357  
; Sequence 34357, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34357  
; LENGTH: 1660  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-34357

Query Match 0.6%; Score 25; DB 6; Length 1660;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTTTTTTTTTTTG 3879  
|||||  
Db 692 AGACTAACTTTTTTTTTTTTTTTG 716

## RESULT 15

US-09-925-065A-34358  
; Sequence 34358, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096

```
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34359
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-34358

Query Match          0.6%; Score 25; DB 6; Length 1660;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTCCTTTTTCG 3879
    |||||
Db 692 AGACTAACTTTTTCCTTTTTCG 716

RESULT 16
US-09-925-065A-34359
; Sequence 34359, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34359
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-34359

Query Match          0.6%; Score 25; DB 6; Length 1660;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3855 AGACTAACTTTTTCCTTTTTCG 3879
    |||||
Db 692 AGACTAACTTTTTCCTTTTTCG 716

RESULT 17
US-11-112-908-63
; Sequence 63, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
```

```
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; LENGTH: 149111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-63

Query Match          0.6%; Score 25; DB 12; Length 149111;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTTTCG 3887
    |||||
Db 148340 TTTTTCCTTTTTCCTTTTTCG 148364

RESULT 18
US-11-121-086-81/c
; Sequence 81, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81
; LENGTH: 156544
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-81

Query Match          0.6%; Score 25; DB 12; Length 156544;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTTTCG 3887
    |||||
Db 27199 TTTTTCCTTTTTCCTTTTTCG 27175

RESULT 19
US-11-112-908-64
; Sequence 64, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
```



; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 64  
; LENGTH: 157230  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-64

Query Match 0.6%; Score 25; DB 12; Length 157230;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
Db 95686 TTTTGTGCTTTT 95710

## RESULT 20

US-11-112-908-62  
; Sequence 62, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:

; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/531,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 62  
; LENGTH: 170508  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-62

Query Match 0.6%; Score 25; DB 12; Length 170508;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
Db 79442 TTTTGTGCTTTT 79466

## RESULT 21

US-11-112-908-65  
; Sequence 65, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:

; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30

; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 65  
; LENGTH: 173115  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-65

Query Match 0.6%; Score 25; DB 12; Length 173115;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3887  
Db 31197 TTTTGTGCTTTT 31221

## RESULT 22

US-09-925-065A-412268/c  
; Sequence 412268, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 412268  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-412268

Query Match 0.6%; Score 24; DB 6; Length 489;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3862 CTTTGTGCTTTT 3885  
Db 136 CTTTGTGCTTTT 113

## RESULT 23

US-09-925-065A-270554/c  
; Sequence 270554, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092

```
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270554
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-270554

Query Match          0.6%; Score 24; DB 6; Length 506;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3886
      |||||
Db 268 TTTTGTGCTTTT 245

RESULT 24
US-09-925-065A-620176/c
; Sequence 620176, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 620176
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-620176

Query Match          0.6%; Score 24; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3886
      |||||
Db 245 TTTTGTGCTTTT 222

RESULT 25
US-09-925-065A-424567
; Sequence 424567, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
```

```
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 424567
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-424567

Query Match          0.6%; Score 24; DB 6; Length 548;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3864 TTTTGTGCTTTT 3887
      |||||
Db 116 TTTTGTGCTTTT 139

RESULT 26
US-09-925-065A-889241/c
; Sequence 889241, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 889241
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-889241

Query Match          0.6%; Score 24; DB 6; Length 549;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3864 TTTTGTGCTTTT 3887
      |||||
Db 46 TTTTGTGCTTTT 23

RESULT 27
US-09-925-065A-624935
; Sequence 624935, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
```



; SEQ ID NO 496602  
; LENGTH: 25  
; TYPE: RNA  
; ORGANISM: Human  
US-10-914A-496602

Query Match 0.6%; Score 23; DB 8; Length 25;  
Best Local Similarity 13.0%; Pred. No. 8.2e+02;  
Matches 3; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 3 UUUUUUUUUUUUUUUUUUUUUUUUUUUUU 25

## RESULT 32

US-10-310-914A-859043  
; Sequence 859043, Application US/10310914A  
; Publication No. US2006003322A1  
; GENERAL INFORMATION:  
; APPLICANT: Bentwich, Isaac  
; APPLICANT: Shiller, Kvuzat  
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and  
; TITLE OF INVENTION: uses thereof  
; FILE REFERENCE: 06087.0200.CPUS01  
; CURRENT APPLICATION NUMBER: US/10/310,914A  
; CURRENT FILING DATE: 2002-12-06  
; NUMBER OF SEQ ID NOS: 1388402  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 859043  
; LENGTH: 25  
; TYPE: RNA  
; ORGANISM: Human  
US-10-310-914A-859043

Query Match 0.6%; Score 23; DB 8; Length 25;  
Best Local Similarity 13.0%; Pred. No. 8.2e+02;  
Matches 3; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 3 UUUUUUUUUUUUUUUUUUUUUUUUUUUUU 25

## RESULT 33

US-11-124-367A-12473  
; Sequence 12473, Application US/11124367A  
; Publication No. US20060024700A1  
; GENERAL INFORMATION:  
; APPLICANT: Michele Cargill  
; APPLICANT: Hongjin Huang  
; TITLE OF INVENTION: Genetic Polymorphisms Associated with  
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof  
; FILE REFERENCE: CL001519.ORD  
; CURRENT APPLICATION NUMBER: US/11/124,367A  
; CURRENT FILING DATE: 2005-05-09  
; PRIOR APPLICATION NUMBER: US 60/568,846  
; PRIOR FILING DATE: 2004-05-07  
; PRIOR APPLICATION NUMBER: US 60/582,609  
; PRIOR FILING DATE: 2004-06-25  
; PRIOR APPLICATION NUMBER: US 60/599,554  
; PRIOR FILING DATE: 2004-08-09  
; NUMBER OF SEQ ID NOS: 34460  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12473  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-124-367A-12473

Query Match 0.6%; Score 23; DB 12; Length 201;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 111 TTTTGTGCTTTT 133

## RESULT 34

US-11-124-367A-12474  
; Sequence 12474, Application US/11124367A  
; Publication No. US20060024700A1  
; GENERAL INFORMATION:  
; APPLICANT: Michele Cargill  
; APPLICANT: Hongjin Huang  
; TITLE OF INVENTION: Genetic Polymorphisms Associated with  
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof  
; FILE REFERENCE: CL001519.ORD  
; CURRENT APPLICATION NUMBER: US/11/124,367A  
; CURRENT FILING DATE: 2005-05-09  
; PRIOR APPLICATION NUMBER: US 60/568,846  
; PRIOR FILING DATE: 2004-05-07  
; PRIOR APPLICATION NUMBER: US 60/582,609  
; PRIOR FILING DATE: 2004-06-25  
; PRIOR APPLICATION NUMBER: US 60/599,554  
; PRIOR FILING DATE: 2004-08-09  
; NUMBER OF SEQ ID NOS: 34460  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12474  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-124-367A-12474

Query Match 0.6%; Score 23; DB 12; Length 201;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 107 TTTTGTGCTTTT 129

## RESULT 35

US-09-925-065A-261182/c  
; Sequence 261182, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 261182  
; LENGTH: 376  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-261182

Query Match 0.6%; Score 23; DB 6; Length 376;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 98281  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-98281

Query Match 0.6%; Score 23; DB 6; Length 498;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTT 3885  
|||||  
Db 221 TTTTTCCTTTTTCCTTTT 243

## RESULT 41

US-09-925-065A-98282  
; Sequence 98282, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 98282  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-98282

Query Match 0.6%; Score 23; DB 6; Length 498;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTT 3885  
|||||  
Db 221 TTTTTCCTTTTTCCTTTT 243

## RESULT 42

US-09-925-065A-925330  
; Sequence 925330, Application US/09925065A

; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 925330  
; LENGTH: 503  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-925330

Query Match 0.6%; Score 23; DB 6; Length 503;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTT 3885  
|||||  
Db 91 TTTTTCCTTTTTCCTTTT 113

## RESULT 43

US-09-925-065A-828324/c  
; Sequence 828324, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 828324  
; LENGTH: 522  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-828324

Query Match 0.6%; Score 23; DB 6; Length 522;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTTCCTTTTTCCTTTT 3885  
|||||  
Db 176 TTTTTCCTTTTTCCTTTT 154



RESULT 44  
US-09-925-065A-819272  
; Sequence 819272, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925.065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 819272  
; LENGTH: 523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-819272

Query Match 0.6%; Score 23; DB 6; Length 523;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3863 TTTTTCCTTTTTCCTTTT 3885  
DB 348 TTTTTCCTTTTTCCTTTT 370

RESULT 45  
US-09-925-065A-3410  
; Sequence 3410, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925.065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3410  
; LENGTH: 556  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-3410

Query Match 0.6%; Score 23; DB 6; Length 556;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3857 ACTAACCTTTTTCCTTTT 3879  
TTTTTTCCTTTTTCCTTTT

Db 444 ACTAACCTTTTTCCTTTT 466

RESULT 46  
US-09-925-065A-755600/c  
; Sequence 755600, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925.065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 755600  
; LENGTH: 574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-755600

Query Match 0.6%; Score 23; DB 6; Length 574;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3864 TTTTTCCTTTTTCCTTTT 3886  
DB 127 TTTTTCCTTTTTCCTTTT 105

RESULT 47  
US-09-925-065A-278724  
; Sequence 278724, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925.065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 278724  
; LENGTH: 575  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-278724

Query Match 0.6%; Score 23; DB 6; Length 575;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3865 TTTTGTGCTTTT 3887  
Db 391 TTTTGTGCTTTT 413

## RESULT 48

US-09-925-065A-561677  
; Sequence 561677, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 561677  
; LENGTH: 576  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-561677

Query Match 0.6%; Score 23; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 27 TTTTGTGCTTTT 49

## RESULT 49

US-11-136-527-5990/c  
; Sequence 5990, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5990  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-11-136-527-5990

Query Match 0.6%; Score 23; DB 12; Length 600;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3863 TTTTGTGCTTTT 3885  
Db 600 TTTTGTGCTTTT 578

## RESULT 50

US-09-925-065A-478484/c  
; Sequence 478484, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 478484  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-478484

Query Match 0.6%; Score 23; DB 6; Length 601;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3856 GACTAACTTTT 3878  
Db 310 GACTAACTTTT 288

Search completed: March 7, 2006, 02:39:09  
Job time : 1043 secs